Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2779157 A1 2011/05/05

(21) 2 779 157

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2010/10/27
(87) Date publication PCT/PCT Publication Date: 2011/05/05
(85) Entrée phase nationale/National Entry: 2012/04/26
(86) N° demande PCT/PCT Application No.: FR 2010/052309
(87) N° publication PCT/PCT Publication No.: 2011/051625
(30) Priorité/Priority: 2009/10/29 (FR09 57636)

(51) Cl.Int./Int.Cl. *H04M 9/08* (2006.01)

(71) Demandeur/Applicant:
UNIVERSITE PARIS DESCARTES, FR

(72) Inventeurs/Inventors:
MAHE, GAEL, FR;
MEZGHANI-MARRAKCHI, IMEN, TN;
JAIDANE, MERIEM, TN;
TURKI, MONIA, TN;
DJAZIRI LARBI, SONIA, TN

(74) Agent: NORTON ROSE CANADA S.E.N.C.R.L.,S.R.L./LLP

(54) Titre : PROCEDE ET DISPOSITIF D'ANNULATION D'ECHO ACOUSTIQUE PAR TATOUAGE AUDIO
(54) Title: METHOD AND DEVICE FOR CANCELLING ACOUSTIC ECHO BY AUDIO WATERMARKING

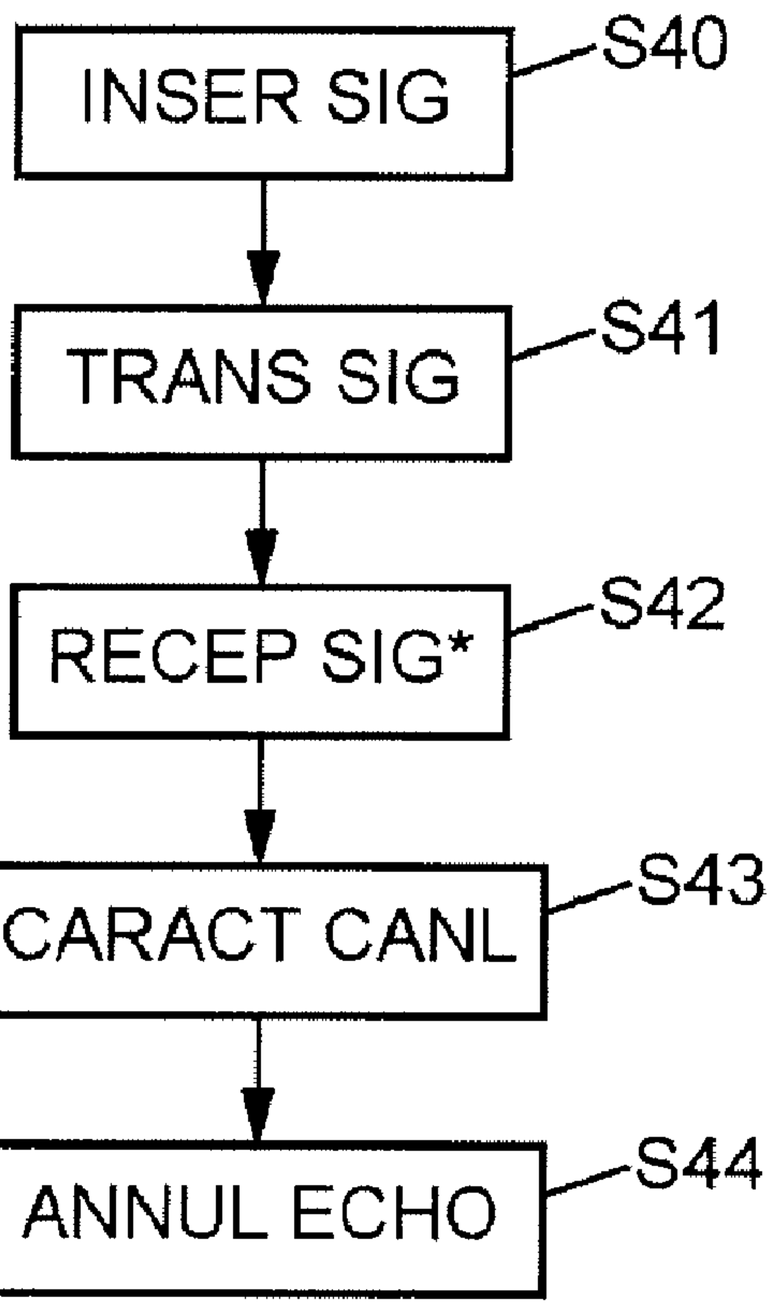


FIG. 4

(57) Abrégé/Abstract:
Procédé d'annulation d'écho acoustique dans un premier signal comportant un signal d'écho d'un deuxième signal, comportant; insérer, de manière inaudible, dans le deuxième signal une séquence pseudo aléatoire dont l'autocorrélation circulaire comporte

CA 2779157 A1 2011/05/05

(21) 2 779 157

(13) A1

(57) Abrégé(suite)/Abstract(continued):
une impulsion unité et une composante continue, caractériser, dans le premier signal, au moyen de la séquence insérée, un canal acoustique suivi par le signal d'écho, estimer le signal d'écho dans le premier signal au moyen de la caractérisation du canal acoustique, et annuler le signal d'écho au moyen de l'estimation obtenue.

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

(43) Date de la publication internationale
5 mai 2011 (05.05.2011)

PCT

(10) Numéro de publication internationale
WO 2011/051625 A1

(51) Classification internationale des brevets :
H04M 9/08 (2006.01)

(21) Numéro de la demande internationale :
PCT/FR2010/052309

(22) Date de dépôt international :
27 octobre 2010 (27.10.2010)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
09 57636 29 octobre 2009 (29.10.2009) FR

(71) Déposant *(pour tous les États désignés sauf US)* : **UNIVERSITE PARIS DESCARTES** [FR/FR]; 12, rue de l'Ecole de Médecine, F-75006 Paris 6 (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants *(pour US seulement)* : **MAHÉ, Gaël** [FR/FR]; 29, rue Carnot, F-78000 Versailles (FR). **MEZGHANI-MARRAKCHI, Imen** [TN/TN]; 34, avenue d'Algérie, Sfax, 3000 (TN). **JAÏDANE, Meriem** [TN/TN]; 3, rue Malaga, El Manar 1, Tunis, 2092 (TN). **TURKI, Monia** [TN/TN]; 7, rue Moncef Bey, Sidi Daoud, 2046 (TN). **DJAZIRI LARBI, Sonia** [TN/TN]; 70-72, rue Medinat Moussol, Cité Ennasr 1, Tunis, 2037 (TN).

(74) Mandataires : **LOISEL, Bertrand** et al.; Cabinet Plasseraud, 52 rue de la Victoire, F-75440 Paris Cedex 09 (FR).

(81) États désignés *(sauf indication contraire, pour tout titre de protection nationale disponible)* : AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Suite sur la page suivante]

(54) Title : METHOD AND DEVICE FOR CANCELLING ACOUSTIC ECHO BY AUDIO WATERMARKING

(54) Titre : PROCEDE ET DISPOSITIF D'ANNULATION D'ECHO ACOUSTIQUE PAR TATOUAGE AUDIO

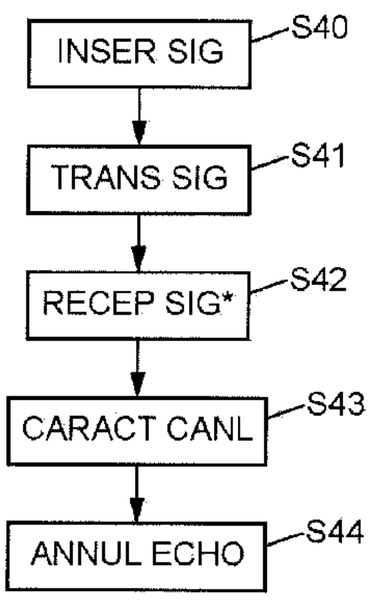


FIG. 4

(57) Abstract : Method of cancelling acoustic echo in a first signal comprising an echo signal of a second signal, comprising; inserting, in an inaudible manner, into the second signal a pseudo-random sequence whose circular autocorrelation comprises a unit impulse and a continuous component; characterizing, in the first signal, by means of the inserted sequence, an acoustic channel followed by the echo signal; estimating the echo signal in the first signal by means of the characterization of the acoustic channel; and cancelling the echo signal by means of the estimation obtained.

(57) Abrégé : Procédé d'annulation d'écho acoustique dans un premier signal comportant un signal d'écho d'un deuxième signal, comportant; insérer, de manière inaudible, dans le deuxième signal une séquence pseudo aléatoire dont l'autocorrélation circulaire comporte une impulsion unité et une composante continue, caractériser, dans le premier signal, au moyen de la séquence insérée, un canal acoustique suivi par le signal d'écho, estimer le signal d'écho dans le premier signal au moyen de la caractérisation du canal acoustique, et annuler le signal d'écho au moyen de l'estimation obtenue.

S40... INSERT SIG

S41... TRANS SIG

S42... RECEP SIG

S43... CHARACTERIZE CHANNEL

S44... CANCEL ECHO

WO 2011/051625 A1

WO 2011/051625 A1

(84) **États désignés** *(sauf indication contraire, pour tout titre de protection régionale disponible)* : ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Déclarations en vertu de la règle 4.17 :

— *relative à la qualité d'inventeur (règle 4.17.iv))*

Publiée :

— *avec rapport de recherche internationale (Art. 21(3))*

Procédé et dispositif d'annulation d'écho acoustique par tatouage audio

Domaine technique

La présente invention concerne le traitement de signaux audio.

Plus particulièrement, elle vise l'annulation de l'écho acoustique dans les systèmes de prise et de restitution du son, notamment utilisés dans des systèmes de communication.

Arrière-plan technologique

Dans les systèmes de prise et de restitution du son de type « mains libres » (comme par exemple les dispositifs de visioconférence), un premier terminal transmet un signal de parole vers un deuxième terminal. Un haut parleur du deuxième terminal émet alors le signal. Le signal émis suit alors un ou plusieurs trajets d'écho. Un microphone du deuxième terminal capte ensuite un signal sonore comportant l'écho du signal émis (et correspondant au signal transmis).

L'écho est réduit par un système dit d'annulation d'écho acoustique (AEC). Dans ces systèmes, le canal acoustique entre le système de restitution et le système de prise de son, comportant un ou plusieurs trajets d'écho, est déterminé par estimation de la réponse impulsionnelle de ce canal acoustique. L'estimation de la réponse impulsionnelle du canal acoustique se fait le plus souvent en temps réel et de manière adaptative auprès du deuxième terminal. Le filtre numérique utilisé pour l'adaptation est piloté par le signal de parole transmis par le premier terminal.

Les performances de tels systèmes sont dégradées par plusieurs facteurs, notamment :

- les caractéristiques statistiques du signal audio (par exemple sa forte corrélation et son caractère non stationnaire, et

- le contexte de la double parole, dans le cas où le deuxième terminal capte un signal de parole locale qui s'ajoute à l'écho et au bruit ambiant. En effet, la présence d'une parole locale s'ajoutant au bruit ambiant gêne la convergence de l'AEC.

Dans les systèmes existants d'annulation d'écho acoustique, qui se fondent sur des méthodes adaptatives, le problème d'existence d'une parole locale se pose du fait que le filtre adaptatif ne peut plus suivre les variations du canal acoustique et suit alors celles de la parole locale. Dans ce cas, l'algorithme converge vers une fausse solution.

Pour remédier à ce problème, l'art antérieur propose de geler l'adaptation en présence de double parole. Cependant, cette solution empêche d'estimer correctement l'écho.

En référence à la Figure 1, on décrit un système d'annulation d'écho (AEC) selon l'art antérieur, mis en œuvre dans un système de communication sonore (par exemple un téléphone en mode « mains libres »).

Un premier interlocuteur envoie, à partir d'un premier terminal, un signal x(n) qui est ensuite diffusé par le haut-parleur HP d'un deuxième terminal. Le signal émis par le haut-parleur est ensuite réfléchi par l'environnement (les murs par exemple) dans lequel le haut-parleur se trouve. La réflexion, ainsi que l'écho direct induit par le passage direct du son entre le haut-parleur et le microphone MIC du deuxième terminal, induisent ainsi un signal d'écho z(n).

Le signal d'écho, un signal de bruit ambiant n(n), ainsi que la parole d'un deuxième interlocuteur s(n) sont captés par le microphone MIC, présent pour capter la parole du deuxième interlocuteur, puis transmis vers le terminal.

Le signal reçu par le premier haut parleur est alors un mélange de la parole du deuxième interlocuteur et de la parole du premier interlocuteur x(n) filtrée par le canal acoustique f entre le haut-parleur et le micro (z(n)=f*x(n), où * désigne la convolution).

Le signal capté par le microphone est alors: y(n)=z(n)+b(n)=(f *x(n))+b(n),

où x(n) et f sont respectivement le signal reçu par le haut parleur de réception et la réponse impulsionnelle du couplage, supposée de nature linéaire, entre le haut-

parleur et le microphone, et b(n)=n(n)+s(n) est un mélange du bruit ambiant n(n) et de la parole locale s(n).

Un système d'annulation d'écho est disposé entre les lignes de transmission des signaux x(n) et y(n), par exemple auprès du deuxième terminal.

Ce système d'annulation d'écho est chargé d'estimer et de simuler au moyen d'un filtre RIF h le canal acoustique f, de manière à retrancher au mélange b(n)+f*x(n) une estimation de f*x(n), donnée par h(n)*x(n).

Le canal de propagation acoustique étant amené à varier dans le temps, des algorithmes adaptatifs sont utilisés.

Leur rôle est "d'apprendre" dans un premier temps puis de mettre continuellement à jour les coefficients du filtre RIF h(n) pour conserver une bonne compensation de l'écho malgré ces variations.

Cependant ces algorithmes adaptatifs souffrent de problèmes de robustesse dus aux variations temporelles des caractéristiques statistiques du signal sonore émis x(n) et pilotant l'AEC (corrélation, non stationnarité,...).

Les coefficients du filtre $h(n) = [h_0(n), h_1(n) \ldots h_{P-1}(n)]^T$ sont par exemple adaptés selon un algorithme normalisé du type NLMS (« Normalized Least Mean Square »).

Cet algorithme minimise l'erreur quadratique moyenne entre la sortie du filtre h(n) et l'écho. L'adaptation de h(n) se fait alors comme suit :

$$h(n+1) = h(n) + \mu e(n) \frac{X(n)}{X(n)^T X(n) + \varepsilon}$$

où μ est un pas d'adaptation généralement fixe, e(n) est l'erreur d'estimation qui contrôle l'adaptation du filtre h(n) et $X(n)=[x(n),x(n-1),\ldots,x(n-P+1)]^T$ est le vecteur d'entrée de longueur P, et ε est une faible valeur strictement positive pour éviter l'annulation du dénominateur de la fraction durant les périodes de silence.

Il a été proposé d'insérer dans le signal incident x(n) un tatouage additif afin de modifier les propriétés statistiques du signal (S. Larbi and M. Jaïdane, "Audio

watermarking: a way to modify audio statistics“, IEEE Trans. on Signal Processing, vol. 53(2), 2005).

Néanmoins, cette méthode présente au moins deux inconvénients :

- l'AEC atténue trop faiblement l'écho et reste sensible aux statistiques du signal de parole tatoué, et

- le problème de la double parole n'est pas résolu.

Il a également été proposé par les inventeurs (Marrakchi et al. « Speech processing in the watermarked domain : application in adaptative acoustic echo cancellation », European Signal Processing Conference, Italy, 2006) un système d'annulation d'écho dit « WAEC » qui est fondé sur l'insertion d'un bruit blanc par tatouage audio dans le signal de parole source de manière à estimer le canal acoustique par un signal ayant de meilleures propriétés que le signal de parole source.

Dans ce système, deux filtres adaptatifs fonctionnent en parallèle.

La nette supériorité du WAEC par rapport à un AEC classique est due au fait que les signaux impliqués dans le fonctionnement du WAEC sont très peu corrélés et beaucoup plus stationnaires. Avec ce système les inventeurs ont obtenu un gain important dans l'atténuation de l'écho.

Toutefois, cette structure peut être améliorée, notamment en présence de double parole.

Résumé de l'invention

Il existe donc un besoin pour un système d'annulation d'écho acoustique offrant une qualité d'annulation d'écho satisfaisante et qui soit peu sensible aux caractéristiques du signal de parole et au contexte de la double parole.

La présente invention vient améliorer la situation.

A cet effet, selon un premier aspect de l'invention, on propose un procédé d'annulation d'écho acoustique dans un premier signal comportant un signal d'écho d'un deuxième signal. Le procédé comporte :

- insérer, de manière inaudible, dans le deuxième signal une séquence pseudo aléatoire dont l'auto corrélation circulaire comporte une impulsion unité et une composante continue,

- caractériser, dans le premier signal, au moyen de la séquence insérée, un canal acoustique suivi par le signal d'écho,

- estimer le signal d'écho dans le premier signal au moyen de la caractérisation du canal acoustique, et

- annuler le signal d'écho au moyen de l'estimation obtenue.

La séquence pseudo aléatoire insérée est indépendante du signal de parole locale ce qui améliore la caractérisation (ou identification) du canal (ou chemin) d'écho.

Ainsi, la qualité de caractérisation est constante et peu dépendante des variations statistiques des signaux mis en œuvre dans l'annulation d'écho.

Avec le procédé selon la présente invention, la qualité de caractérisation est meilleure, même en contexte de double parole.

La présente invention permet de fournir une qualité satisfaisante de conversation téléphonique ou en visioconférence, même en présence des variations statistiques de la parole émise et de la parole locale.

En outre, avec le procédé selon la présente invention, l'annulation d'écho est plus rapide et plus précise que dans l'art antérieur.

Par exemple, la caractérisation du canal acoustique (ou chemin d'écho) se fait par intercorrélation du premier signal avec la séquence pseudo aléatoire.

Cela permet de minimiser l'effet de l'existence d'une parole locale au cours de l'identification à cause de l'indépendance statistique de la séquence pseudo aléatoire et de la parole locale.

Par exemple encore, la caractérisation du canal acoustique comporte un filtrage adaptatif du deuxième signal.

La caractérisation peut également comporter un filtrage par bloc du deuxième signal.

L'utilisation de la séquence pseudo aléatoire permet de réduire la perturbation du signal en présence de double parole.

Dans des modes de réalisation, la caractérisation du canal acoustique comporte les étapes :

- effectuer un filtrage adaptatif et/ou par bloc du deuxième signal, et

- effectuer une intercorrélation du premier signal avec la séquence pseudo aléatoire.

Le filtrage permet d'éliminer l'influence du deuxième signal avant l'intercorrélation, augmentant ainsi la qualité de cette dernière. Par exemple, il est possible de diminuer la puissance du deuxième signal par rapport à la puissance de la séquence pseudo aléatoire.

Dans des modes de réalisation, le procédé comporte en outre :

- appliquer un filtre de mise en forme à la séquence pseudo aléatoire avant son insertion dans le deuxième signal.

Par exemple, le filtre de mise en forme est utilisé pour assurer l'inaudibilité de la séquence pseudo aléatoire, en se fondant sur des modèles psychoacoustiques.

Dans ce cas, on peut former le filtre à fournir une séquence inaudible avec une puissance maximale, pour optimiser son utilisation dans l'annulation d'écho.

Dans des modes de réalisation, on peut prévoir en outre une étape de traitement du premier signal avant la corrélation avec la séquence pseudo aléatoire. Ce traitement peut comporter la soustraction du résultat du filtrage du deuxième signal au premier signal. Ainsi, la qualité du résultat de la corrélation peut être améliorée.

Par exemple, le traitement peut aussi comporter l'application d' un filtre permettant d'annuler les effets du filtre de mise en forme sur la séquence pseudo aléatoire.

Selon d'autres aspects de l'invention, il est prévu :

- un programme informatique comportant des instructions pour la mise en œuvre du procédé selon l'invention lorsque le programme est exécuté par un processeur, par exemple, le processeur d'un système d'annulation d'écho ;

- un support lisible par un ordinateur sur lequel est enregistré un tel programme d'ordinateur ;

- un circuit configuré pour la mise en œuvre du procédé selon l'invention ; et

- un système d'annulation d'écho adapté pour la mise en œuvre du procédé selon l'invention.

Les avantages procurés par le programme informatique, le support lisible par ordinateur, le circuit, et le système, tels que succinctement exposés ci-dessus, sont au moins identiques à ceux mentionnés plus haut en liaison avec le procédé de d'annulation d'écho selon l'invention.

Brève description des dessins

D'autres caractéristiques et avantages de l'invention apparaîtront encore à la lecture de la description qui va suivre. Celle-ci est purement illustrative et doit être lue en regard des dessins annexés sur lesquels, en outre la Figure 1 :

- la figure 2 est un schéma bloc illustrant un procédé selon un premier mode de réalisation de l'invention ;

- la figure 3 est un schéma bloc illustrant un procédé selon un deuxième mode de réalisation de l'invention ;

- la figure 4 est un organigramme général du procédé selon des modes de réalisation de l'invention ;

- la figure 5 illustre un système d'annulation d'écho selon un mode de réalisation de l'invention ; et

- les figures 6 à 9 illustrent des résultats de tests montrant quelques avantages apportés par l'invention.

Description détaillée de modes de réalisation

En référence à la Figure 2, on décrit schématiquement la mise en œuvre d'un procédé selon un premier mode de réalisation.

Dans ce mode de réalisation, l'annulation d'écho se fait par insertion d'une séquence numérique (procédé également appelé « tatouage numérique »), puis par

intercorrélation de la séquence insérée avec le signal comportant le signal d'écho pour la caractérisation (ou identification) du canal acoustique (ou chemin d'écho). L'annulation finale de l'écho se fait par soustraction, dans le signal comportant l'écho, d'une estimation de l'écho fondée sur l'identification du canal acoustique.

Dans un premier temps, on décrit de manière générale les étapes mises en œuvre. Dans un second temps, chacune des étapes sera détaillée.

Ce mode de réalisation s'inscrit par exemple dans le cadre d'une communication entre un premier terminal de communication tel qu'un téléphone mobile, et un deuxième terminal en mode mains libres.

L'écho que l'on se propose de traiter prend naissance auprès du deuxième terminal, et est traité auprès de ce deuxième terminal.

On peut envisager d'autres modes de réalisation dans lesquels l'écho est traité auprès du premier terminal, ou encore dans lesquels l'écho est traité dans un réseau de télécommunications.

Le signal de parole x initialement émis par un utilisateur du premier terminal est fourni à un bloc d'insertion INS, pour l'insertion d'une séquence numérique qui sera décrite plus en détails dans la suite.

Le signal inséré t provient d'un signal initial w mis en forme par un bloc de mise en forme FORM. Cette mise en forme permet d'adapter le spectre du signal w afin de le rendre inaudible, une fois inséré dans le signal x.

Le signal xt obtenu à la sortie du bloc INS est ensuite transmis et suit un ensemble de transformations modélisé par le bloc ECHO. Ces transformations correspondent par exemple à sa transmission et à sa réflexion.

A la sortie du bloc ECHO on obtient un signal z, comme celui décrit précédemment en référence à la Figure 1.

A ce signal z, un signal de bruit b vient également s'ajouter, pour former le signal y. Le signal y contient l'écho du signal x.

Le signal y est ensuite fourni à une unité de mise en forme du signal DEFORM qui effectue sur le signal une transformation inverse de celle effectuée par le bloc FORM.

En sortie du bloc DEFORM on obtient un signal yf qui est ensuite fourni à un bloc d'intercorrélation INTERCOR afin de réaliser l'intercorrélation du signal yf avec le signal w.

Le bloc INTERCOR fournit alors une caractérisation $\hat{f}$ du canal acoustique schématisé par le bloc ECHO. Cette caractérisation est ensuite fournie à un bloc ESTIM qui reçoit également en entrée le signal xt, pour obtenir en sortie une estimation du signal z.

Le signal y et l'estimation du signal z sont ensuite fournis à un bloc ANNUL pour soustraire l'écho du signal y.

Dans un deuxième mode de réalisation, décrit en référence à la Figure 3, un filtrage adaptatif est effectué sur le signal tatoué pour réaliser une première identification du canal acoustique, avant l'intercorrélation.

Dans ce mode de réalisation, on retrouve les blocs FORM, INS, ECHO, ANNUL, INTERCOR, et ESTIM du premier mode de réalisation.

En outre, on trouve un deuxième bloc DEFORM identique au premier, et alimenté par le signal xt, la sortie de ce deuxième bloc DEFORM est couplée à l'entrée d'un bloc ADAPT1 de filtrage adaptatif.

Le signal w est par ailleurs fourni à un deuxième bloc ADAPT2, qui est une copie du premier bloc ADAPT1.

Dans ce deuxième mode de réalisation, la sortie du premier bloc DEFORM est fournie à un bloc de soustraction pour effectuer la soustraction entre le signal yf et la sortie du premier bloc ADAPT1. Le résultat de cette soustraction est délivré au bloc ADAPT1 pour piloter le filtrage.

La sortie du deuxième bloc ADAPT2 est ensuite additionnée avec la sortie du bloc de soustraction, et la somme est fournie au bloc INTERCOR.

Dans une variante du deuxième mode de réalisation (non représentée), on utilise un filtrage par bloc alternativement au, ou en combinaison avec le, filtrage adaptatif. Par exemple, on implémente un filtre de Wiener.

Les étapes du procédé selon les modes de réalisation décrits, sont résumées dans l'organigramme général de la Figure 4.

Dans une première étape S40, le signal transmis par l'utilisateur du premier terminal est reçu par le deuxième terminal puis est tatoué par l'AEC par insertion de la séquence numérique. Ensuite, le signal tatoué est diffusé par le haut parleur du deuxième terminal lors de l'étape S41.

Lors de l'étape S42, le microphone du deuxième terminal reçoit un signal comportant un signal d'écho du signal précédemment transmis.

On réalise alors lors de l'étape S43 une caractérisation du canal acoustique au moyen de la séquence insérée lors du tatouage.

Enfin, lors de l'étape S44 l'écho est annulé dans le signal reçu par le microphone, au moyen de la caractérisation obtenue.

Un programme d'ordinateur comportant des instructions pour la mise en œuvre du procédé selon l'invention peut être réalisé selon un algorithme général déduit de l'organigramme général de la Figure 4, et de la présente description détaillée.

Un système d'annulation d'écho selon un mode de réalisation de l'invention est décrit en référence à la Figure 5.

Ce système d'annulation d'écho AEC comporte une entrée E2 pour recevoir un signal à retransmettre, une sortie S2 pour retransmettre le signal reçu et tatoué. Il comporte également une entrée E1 pour recevoir un signal comportant un écho du signal transmis, et une sortie S1 pour envoyer le signal avec l'écho annulé.

Le système AEC comporte également une mémoire MEM pour stocker des données de calcul. Dans des modes de réalisation, la mémoire MEM peut également stocker un programme d'ordinateur selon la présente invention.

Le système comporte également un processeur PROC pour contrôler un circuit d'annulation d'écho. Par exemple, le processeur exécute un programme d'ordinateur stocké dans la mémoire MEM.

Le circuit CIRC comporte une unité d'insertion de séquence numérique INS, une unité de caractérisation de canal acoustique CARACT, une unité d'estimation du signal d'écho ESTIM et une unité d'annulation d'écho ANNUL.

Tous ces éléments sont disposés en sorte à fonctionner selon le procédé d'annulation d'écho selon l'invention.

Le système d'annulation d'écho peut faire partie d'un terminal de communication. Par exemple, il peut faire partie d'un terminal de communication mettant en œuvre une communication en mode mains libres. Ainsi, l'écho naissant auprès de ce terminal peut être directement annulé avant de retransmettre le signal de parole capté. Par exemple encore, le système peut faire partie d'un terminal de communication qui ne met pas en œuvre une communication en mode mains libres, mais qui est en communication avec un terminal mettant en œuvre une telle communication. Ainsi, le terminal peut éliminer l'écho dans un signal reçu.

Par exemple encore, le système d'annulation d'écho est mis en œuvre dans un serveur de communication. Ainsi, un opérateur d'un réseau de télécommunications mettant en œuvre le serveur peut fournir à ses abonnés un service d'annulation d'écho.

L'intégration d'un système selon l'invention dans un terminal ou un serveur, tels que succinctement décrits, peut être réalisée selon des techniques connues de la personne du métier.

Dans la suite, on décrit plus en détails les différentes opérations évoquées ci-avant : le tatouage numérique, la caractérisation du canal acoustique, l'annulation d'écho, et le filtrage adaptatif ou par bloc.

Enfin, on présente des résultats de tests effectués pour montrer certains avantages apportés par la présente invention.

Tatouage numérique

Le signal w(n) inséré est une séquence pseudo aléatoire de +1 et -1 de longueur L, appelée séquence à longueur maximale (MLS), périodisée. Elle a comme principale propriété de donner une impulsion unité L-périodisée plus une composante continue 1/L lorsque l'on effectue son auto-corrélation circulaire :

$$\phi_{ww}(n) = \begin{cases} 1 \text{ si } n = 0 \bmod L \\ -\frac{1}{L} \text{ si } n \neq 0 \end{cases}$$

où l'auto-corrélation circulaire est définie par :

$$\phi_{ww}(n) = \frac{1}{L}\sum_{k=0}^{L-1} w(k)w((n+k) \bmod L)$$

L'insertion temporelle du signal de tatouage dans les signaux audio se fait de manière inaudible en utilisant des modèles psychoacoustiques.

Le seuil de masquage, dans le cas de la parole, est approché par la Densité Spectrale de Puissance (DSP) du signal sur une trame de 20ms, atténuée d'un facteur $\lambda<1$. Comme la parole est modélisable par la filtrée du bruit blanc de variance σ_{ex}^2 par un filtre tout-pôles de fonction de transfert 1/A(z), on peut ajouter un tatouage inaudible obtenu par le filtrage d'un bruit blanc w(n) par le filtre de fonction de transfert $G(z) = \frac{\lambda\sigma_{ex}}{A(z)}$. Ce filtre est actualisé tous les L échantillons.

Caractérisation du canal par intercorrélation

Il s'agit d'exciter le canal par une séquence MLS w(n) de longueur L périodisée. Le signal y(n) obtenu à la sortie du canal est :

$$y(n) = f * w(n) = \sum_{j=0}^{P-1} f(j)w(n-j)$$

où P est la longueur de la réponse impulsionnelle f à estimer.

La corrélation entre la sortie y(n) et l'entrée w(n) vaut :

$$\begin{aligned}\phi_{wy}(n) &= \frac{1}{L}\sum_{k=0}^{L-1} w(k)\,y(n+k) \\ &= \frac{1}{L}\sum_{k=0}^{L-1}\sum_{j=0}^{P-1} f(j)w(k)w(n+k-j) \\ &= \sum_{j=0}^{P-1} f(j)\phi_{ww}(n-j)\end{aligned}$$

Ou encore :

$$\phi_{wy}(n) = \hat{f}(n) = f(n) + P_1(n) + P_2(n)$$

avec

$P_1(n) = \sum_{j=1}^{[P/L]} f(n + jL)$ est la perturbation due à l'effet de sous-modélisation lorsque L<P.

$P_2(n) = -\frac{1}{L} \sum_{j=1, j \neq n+kL}^{P-1} f(j)$ est la perturbation due à l'effet de la « «fausse impulsion unité».

La situation idéale est réalisée dans le cas de modélisation exacte (P<L) où $P_1(n) = 0$ et pour L suffisamment grand et dans ce cas $P_2(n) \approx 0$.

Annulation d'écho

Le signal de tatouage obtenu est ajouté temporellement au signal de parole comme illustré par la Figure 2. Le signal tatoué xt(n) obtenu est ensuite transmis vers le canal acoustique f à identifier. L'écho obtenu est donné par :

$$\begin{aligned} y(n) &= f * xt(n) + b(n) \\ &= f * x(n) + f * t(n) + b(n) \\ &= f * x(n) + f * g * w(n) + b(n) \end{aligned}$$

En appliquant au signal d'écho l'inverse du filtre de mise en forme g du bloc FORM, on obtient :

$$\begin{aligned} yf(n) &= g^{-1} * y(n) \\ &= f * w(n) + f * xf(n) + bf(n) \end{aligned}$$

où xf(n) et bf(n) sont respectivement les signaux x(n) et b(n) filtrés par le filtre g^{-1}.

L'estimation du canal se fait par calcul d'intercorrélation par bloc de L échantillons entre le signal d'écho filtré et la séquence MLS originale w(n) :

$$\phi_{wyf}(n) = f(n) + f * \phi_{wxf}(n) + \phi_{wbf} + P_1(n) + P_2(n), n = 0 : L-1$$

L'écho résiduel est donc :

$$e(n) = y(n) - \phi_{wyf}(n) * xt(n)$$

La qualité de l'estimation est indépendante de la corrélation et de la non stationnarité du signal de parole émis.

Identification adaptative ou par bloc complémentaire

Afin d'assurer la rapidité de la caractérisation du canal acoustique, la valeur de L doit être limitée.

Dans ce cas, l'intercorrélation ϕ_{wxf} n'est pas parfaitement nulle et par conséquent le terme $f * \phi_{wxf}(n)$, qui est de puissance un peu plus faible que celle de $\phi_{wxf}(n)$, ne peut être totalement négligé.

Afin de pallier ce problème, on peut prévoir l'annulation de l'effet de ce terme par l'ajout d'une étape de filtrage adaptatif.

Ainsi, on obtient une première estimation du canal acoustique f comme illustré par la Figure 3.

L'étage de filtrage adaptatif ADAPT (filtre adaptatif h(n) de longueur P) est piloté par le signal tatoué filtré $xtf(n) = g^{-1} * xt(n)$ et contrôlé par l'erreur d'estimation adaptative :

$$\begin{aligned} er(n) &= yf(n) - h(n) * [xf(n) + w(n)] \\ &= f * w(n) + \underbrace{v(n) * xf(n) + bf(n)}_{\xi(n)} - h(n) * w(n) \end{aligned}$$

où v(n)=f-h(n) est le vecteur déviation qui représente l'erreur d'estimation du canal.

Soit le signal à l'entrée du corrélateur INTERCOR :

$$\begin{aligned} e(n) &= er(n) + h(n) * w(n) \\ &= f * w(n) + \xi(n) \end{aligned}$$

A la convergence du filtre adaptatif h(n), la puissance de l'erreur $\xi(n)$ converge vers celle du bruit filtré bf(n).

L'estimation du canal se fait par calcul de l'intercorrélation :

$$\phi_{we}(n) = \hat{f}(n) + \phi_{w\xi}(n), n = 0 : L-1$$

Dans le cas où le filtre h(n) converge, le second terme de l'équation devient négligeable et on obtient une qualité d'estimation quasiment constante.

Selon une variante de réalisation, on peut utiliser un filtrage par bloc. Par exemple, il s'agit d'utiliser un filtre de Wiener. Comme il apparaîtra à la personne du métier, l'utilisation d'un tel filtre peut nécessiter d'inverser une matrice de corrélation du signal audio.

Résultats

Différents signaux de parole et différents canaux acoustiques ont été utilisés pour tester le procédé selon l'invention.

Les tests ont montré que la qualité de la parole à la réception est nettement meilleure avec la méthode proposée qu'avec celles utilisant les méthodes adaptatives classiques dans lesquelles l'AEC est piloté par le signal de parole reçu directement.

Avec le procédé selon la présente invention, le temps d'adaptation est plus rapide et une estimation plus stationnaire de l'écho est réalisée.

Pour avoir des résultats comparables, les performances de la méthode proposée, dont l'étage adaptatif est de type NLMS, ont été comparées à celle d'un AEC classique.

L'AEC classique en question est un AEC adaptatif type NLMS dont l'entrée est le signal d'entrée x(n) lui-même pilotant le canal acoustique à estimer.

Le même pas d'adaptation μ a été utilisé pour les deux AEC (AEC classique et AEC selon la présente invention).

Les paramètres de simulation utilisés ont été les suivants : L=511, P=200, en absence de parole locale.

La Figure 6 illustre l'évolution de la déviation quadratique moyenne (DQM), pour un procédé selon l'invention (courbe AEC_INV) et pour un procédé selon l'art antérieur (courbe AEC_CLASS). La DQM représente l'erreur relative d'estimation de f, et s'exprime selon :

$$DQM_{dB}(n) = 10Log\left[\frac{E\left(\left|\phi_{we}(n)\right|^2\right)}{\left|f\right|^2}\right].$$

La DQM est calculée en l'absence de parole locale et en présence de bruit ambiant avec un rapport signal à bruit de 20dB. Le rapport signal à bruit s'exprime selon:

$$RSB1 = 10Log\left[\frac{P_x}{P_n}\right]$$

avec Px la puissance du signal d'entrée x(n) et Pn celle du bruit ambiant n(n).

Comme on peut le constater sur la Figure 6, la vitesse de convergence réalisée par la méthode proposée est nettement supérieure à celle d'un AEC classique.

En outre, la déviation finale est plus faible avec la présente invention.

Le procédé selon la présente invention est donc plus rapide et plus précis.

Afin d'évaluer les performances de l'AEC en régime permanent, on a calculé l'ERLE (« echo return loss enhancement »).

L'ERLE est défini par :

$$ERLE = 10Log_{10}\left[\frac{E\left(\{y(n)\}^2\right)}{E\left(\{e(n)\}^2\right)}\right]$$

avec y(n) l'écho bruité à estimer et e(n) est l'erreur d'estimation.

Les Figures 7 et 8 illustrent, pour un signal de parole SIG initialement émis et donnant naissance à un signal d'écho, l'évolution de l'ERLE respectivement pour la structure proposée (courbe AEC_INV) et pour un AEC classique (par exemple type NLMS, courbe AEC_CLASS) également en absence de parole locale et en présence de bruit ambiant et pour la structure proposée et le WAEC (courbe WAEC) également dans les mêmes conditions.

Cette comparaison en régime permanent montre que la structure proposée garantit une qualité d'estimation meilleure et plus stable que celle offerte par un AEC classique et par le WAEC.

La figure 9 illustre, pour un signal de parole SIG initialement émis et donnant naissance à un signal d'écho, l'évolution de l'ERLE en présence de bruit ambiant et de parole locale SIG_LOC, dans le cas de la présente invention et d'un AEC classique. La puissance de la parole locale est généralement liée à celle de l'écho par le rapport :

$$RSB2 = 10Log\left[\frac{P_{echo}}{P_s}\right]$$

Dans le présent cas, RSB2= -20 dB. Les résultats montrent que la structure proposée donne de meilleures performances même en présence de parole locale.

La présente invention a été décrite et illustrée dans la présente description détaillée et dans les Figures. La présente invention ne se limite pas aux formes de réalisation présentées. D'autres variantes et modes de réalisation peuvent être déduits et mis en œuvre par la personne du métier à la lecture de la présente description et des Figures annexées.

Dans les revendications, le terme "comporter" n'exclut pas d'autres éléments ou d'autres étapes. L'article indéfini « un » n'exclut pas le pluriel. Un seul processeur ou plusieurs autres unités peuvent être utilisées pour mettre en œuvre l'invention. Les différentes caractéristiques présentées et/ou revendiquées peuvent être avantageusement combinées. Leur présence dans la description ou dans des revendications dépendantes différentes, n'excluent pas cette possibilité. Les signes de référence ne sauraient être compris comme limitant la portée de l'invention.

REVENDICATIONS

1. Procédé d'annulation d'écho acoustique dans un premier signal (y) comportant un signal d'écho d'un deuxième signal (x), caractérisé en ce qu'il comporte :

- insérer (S40), de manière inaudible, dans le deuxième signal une séquence pseudo aléatoire (w) dont l'autocorrélation circulaire comporte une impulsion unité et une composante continue,

- caractériser (S43), dans le premier signal, au moyen de la séquence insérée, un canal acoustique suivi par le signal d'écho,

- estimer (ESTIM) le signal d'écho dans le premier signal au moyen de la caractérisation du canal acoustique, et

- annuler (S44) le signal d'écho au moyen de l'estimation obtenue.

2. Procédé selon la revendication 1, dans lequel la caractérisation du canal acoustique se fait par intercorrélation (INTERCOR) du premier signal avec la séquence pseudo aléatoire.

3. Procédé selon la revendication 1, dans lequel la caractérisation du canal acoustique comporte au moins l'un d'un filtrage adaptatif (ADAPT) du deuxième signal et d'un filtrage par bloc du deuxième signal.

4. Procédé selon la revendication 1, dans lequel la caractérisation du canal acoustique comporte les étapes :

- effectuer un filtrage adaptatif et/ou par bloc du deuxième signal, et

- effectuer une intercorrélation du premier signal avec la séquence pseudo aléatoire.

5. Procédé selon l'une des revendications précédentes, comportant en outre :

- appliquer un filtre de mise en forme (FORM) à la séquence pseudo aléatoire avant son insertion dans le deuxième signal.

6. Procédé selon l'une des revendications 3 à 5, comportant en outre une étape de traitement du premier signal avant l'intercorrélation avec la séquence pseudo aléatoire, ledit traitement comportant une soustraction d'un résultat du filtrage du deuxième signal au premier signal.

7. Programme d'ordinateur comportant des instructions pour la mise en œuvre d'un procédé selon l'une des revendications précédentes lorsqu'il est exécuté par un processeur.

8. Circuit configuré pour la mise en œuvre d'un procédé selon l'une des revendications 1 à 6.

9. Système d'annulation d'écho (AEC) dans un premier signal (y) comportant un signal d'écho d'un deuxième signal (x), le système comportant :

- une unité d'insertion (INS), de manière inaudible, dans le deuxième signal d'une séquence pseudo aléatoire dont l'autocorrélation circulaire comporte une impulsion unité et une composante continue,

- une unité de caractérisation (CARACT), dans le premier signal, au moyen de la séquence insérée, d'un canal acoustique suivi par le signal d'écho,

- une unité d'estimation (ESTIM) du signal d'écho dans le premier signal au moyen de la caractérisation du canal acoustique, et

- une unité d'annulation (ANNUL) du signal d'écho au moyen de l'estimation obtenue.

10. Système selon la revendication 9, dans lequel l'unité de caractérisation du canal acoustique comporte une unité d'intercorrélation du premier signal avec la séquence pseudo aléatoire.

11. Système selon la revendication 9, dans lequel l'unité de caractérisation du canal acoustique comporte au moins l'un d'un filtre adaptatif du deuxième signal, et d' un filtre par bloc du deuxième signal.

12. Système selon la revendication 9, dans lequel l'unité de caractérisation du canal acoustique comporte :

- un filtre adaptatif et/ou par bloc du deuxième signal, et
- une unité d'intercorrélation du premier signal avec la séquence pseudo aléatoire.

13. Système selon l'une des revendications 9 à 12, comportant en outre :

- un filtre de mise en forme de la séquence pseudo aléatoire avant son insertion dans le deuxième signal.

14. Système selon l'une des revendications 11 à 13, comportant en outre une unité de traitement du premier signal avant l'intercorrelation avec la séquence pseudo aléatoire, ledit traitement comportant une soustraction d'un résultat du filtrage du deuxième signal au premier signal.

15. Système selon l'une des revendications 9 à 14, comportant en outre :

- une première entrée (E1) pour recevoir le premier signal,
- une première sortie (S1) pour envoyer le premier signal avec le signal d'écho annulé,
- une deuxième entrée (E2) pour recevoir le deuxième signal, et
- une deuxième sortie (S2) pour envoyer le deuxième signal avec la séquence insérée.

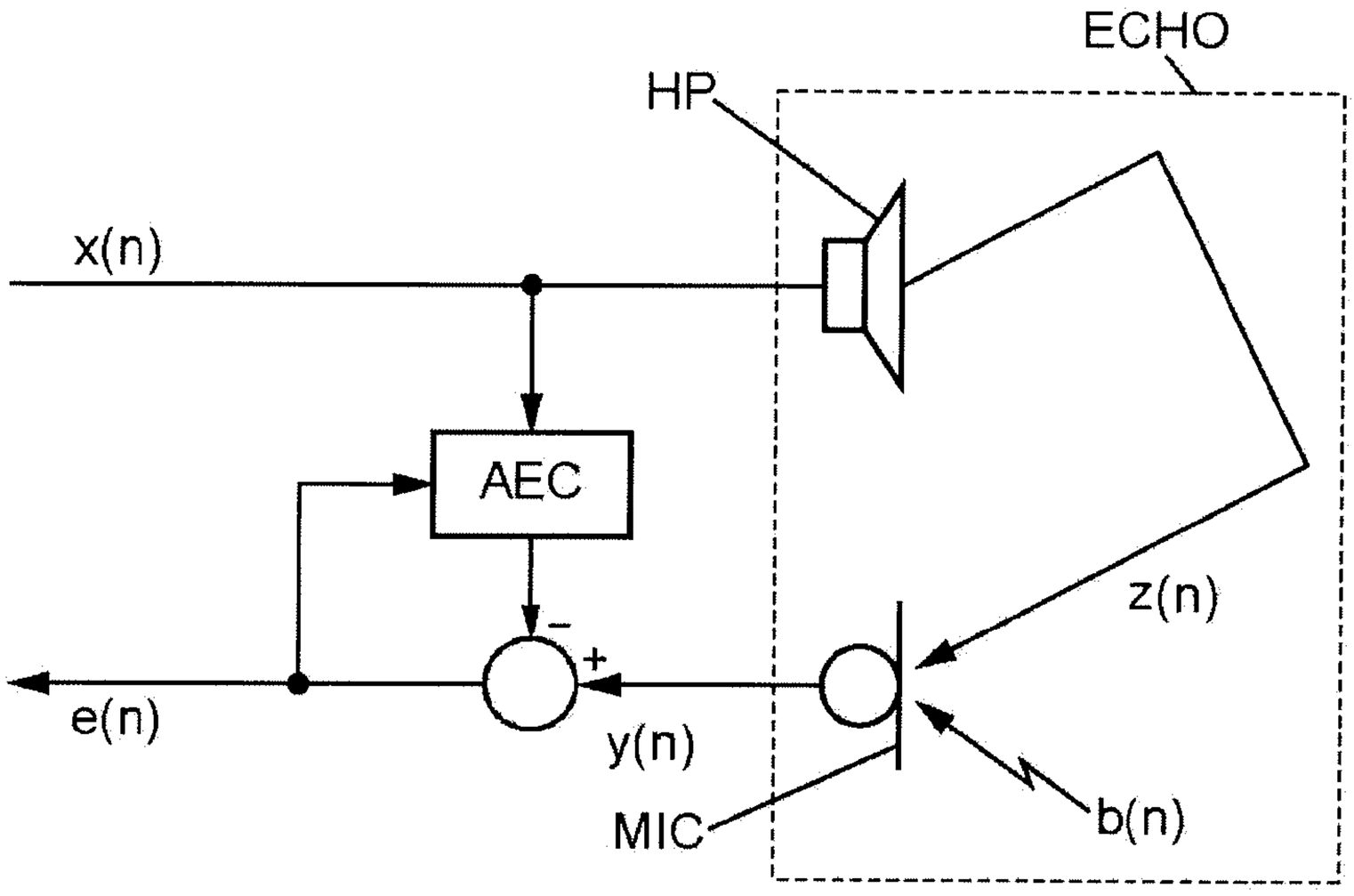

ECHO
HP
x(n)
AEC
z(n)
–
+
e(n)
y(n)
MIC
b(n)

FIG. 1

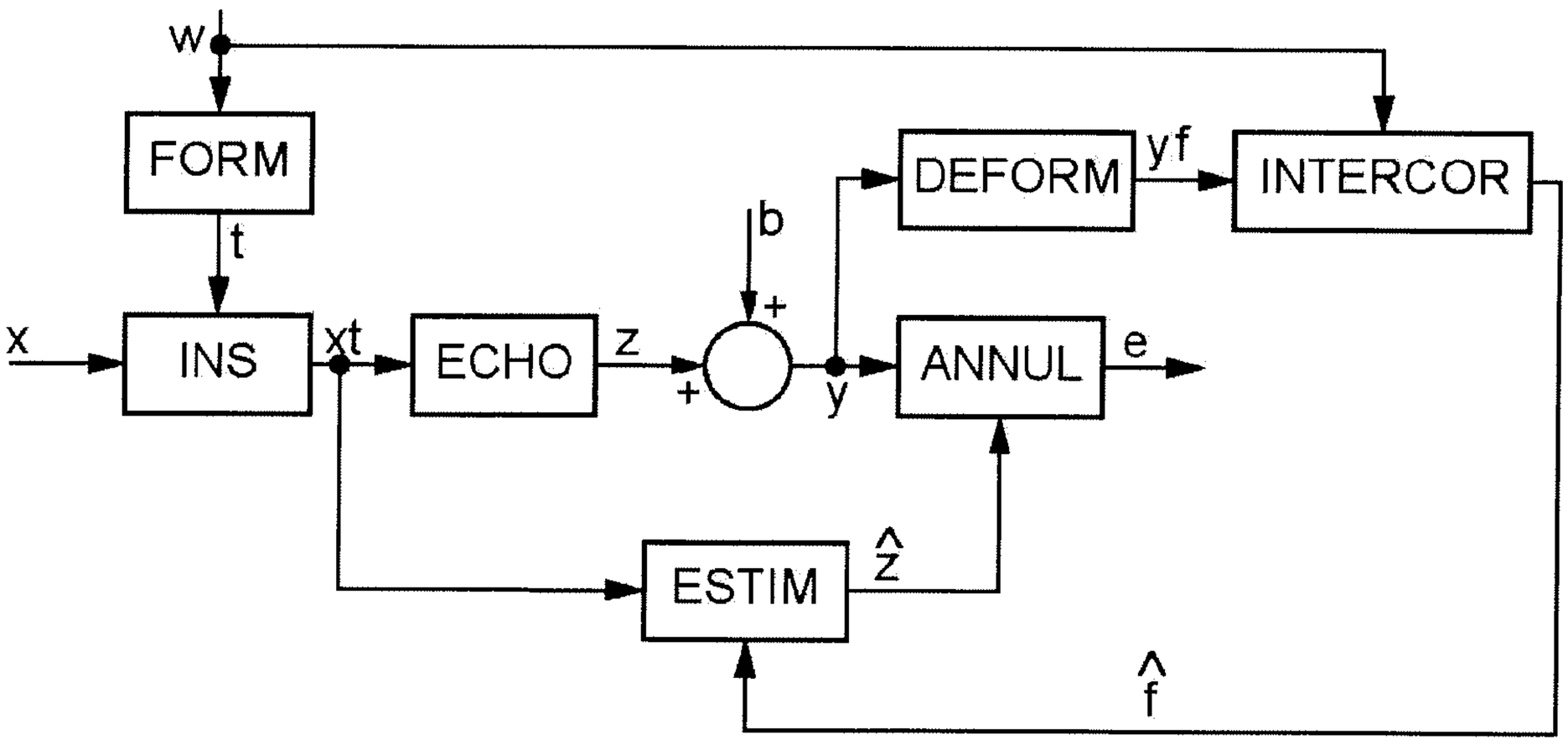

w
FORM
t
x
INS
xt
ECHO
z
b
+
+
y
DEFORM
yf
INTERCOR
ANNUL
e
ESTIM
$\hat{z}$
$\hat{f}$

FIG. 2

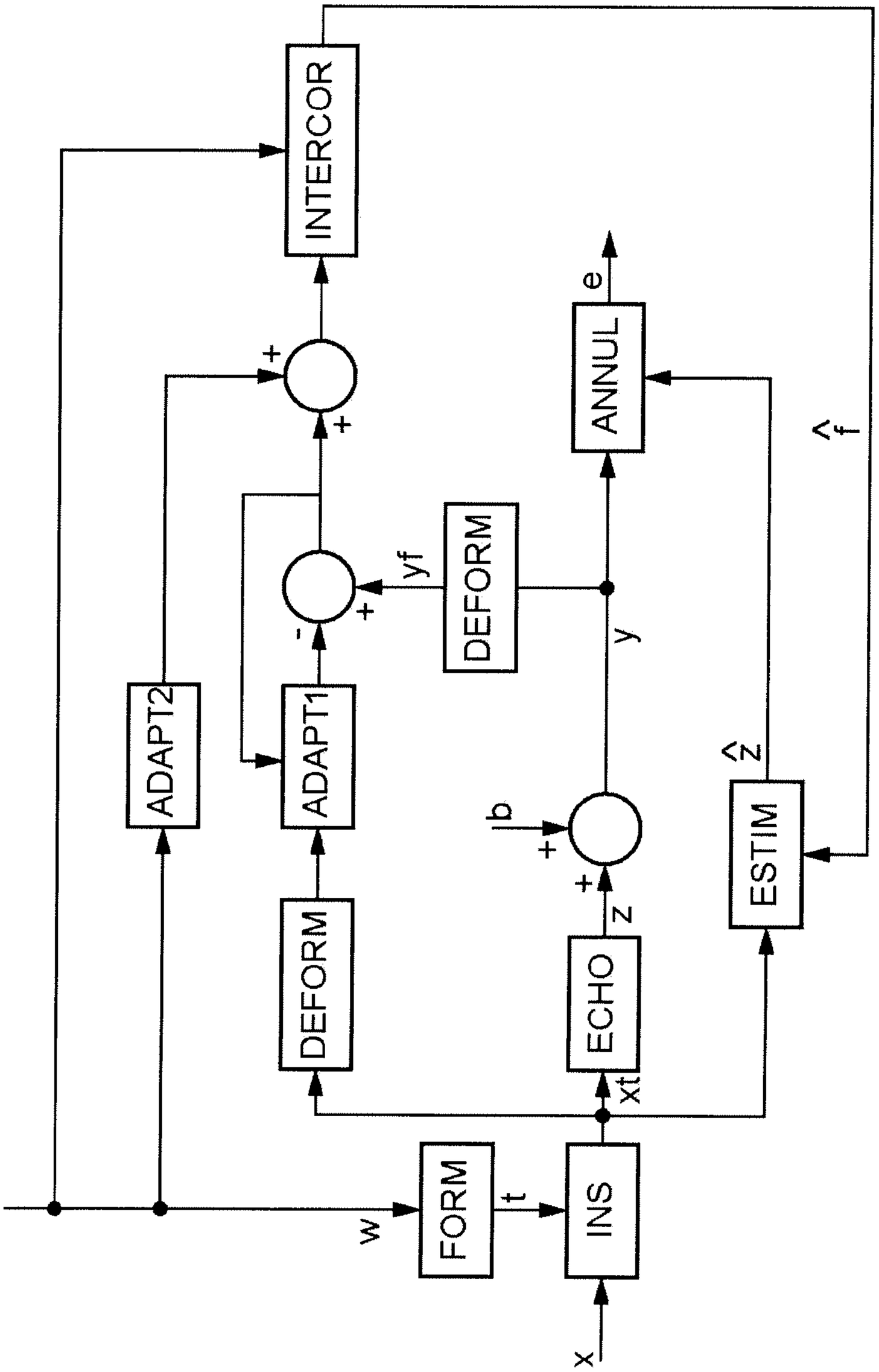

INTERCOR
+
+
ANNUL
e
^
f
yf
DEFORM
+
-
y
ADAPT2
ADAPT1
DEFORM
b
+
+
ECHO
z
^
z
ESTIM
xt
FORM
t
INS
w
x

FIG. 3

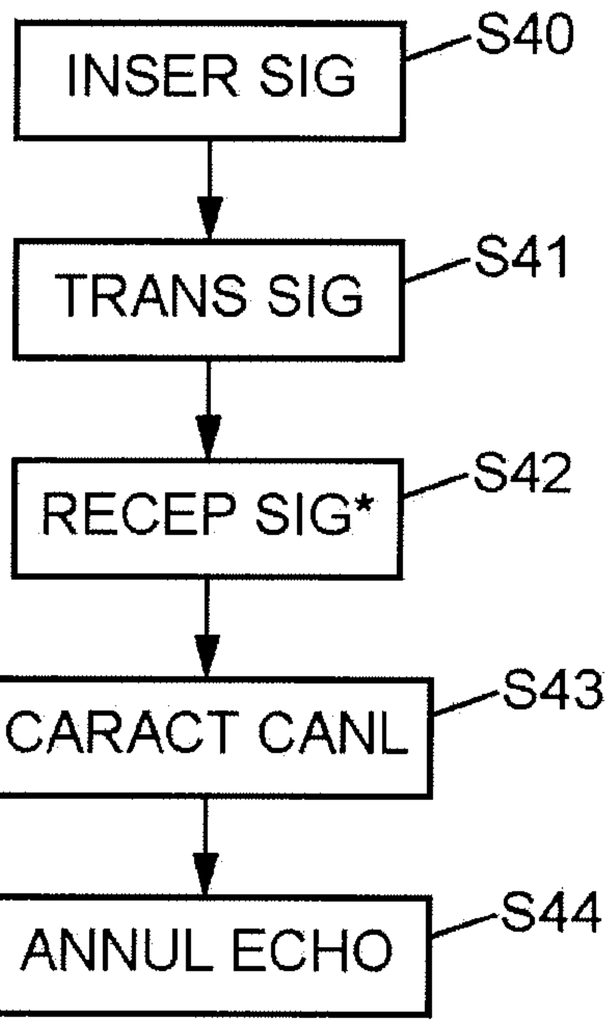

INSER SIG
S40
TRANS SIG
S41
RECEP SIG*
S42
CARACT CANL
S43
ANNUL ECHO
S44

FIG. 4

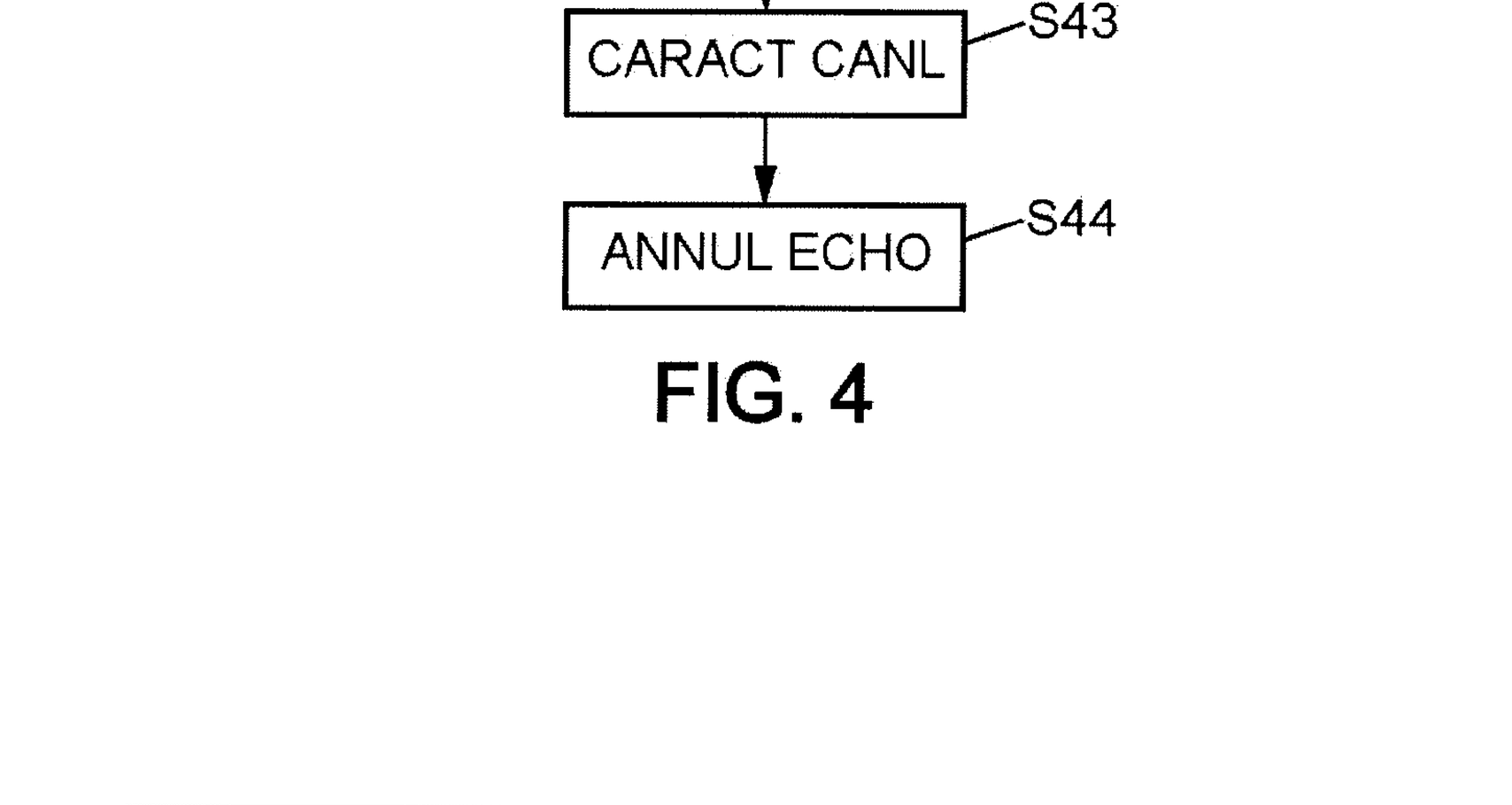

AEC
MEM
PROC
CIRC
x
E2
INS
S2
xt
e
S1
ANNUL
ESTIM
CARACT
E1
y

FIG. 5

FIG. 6

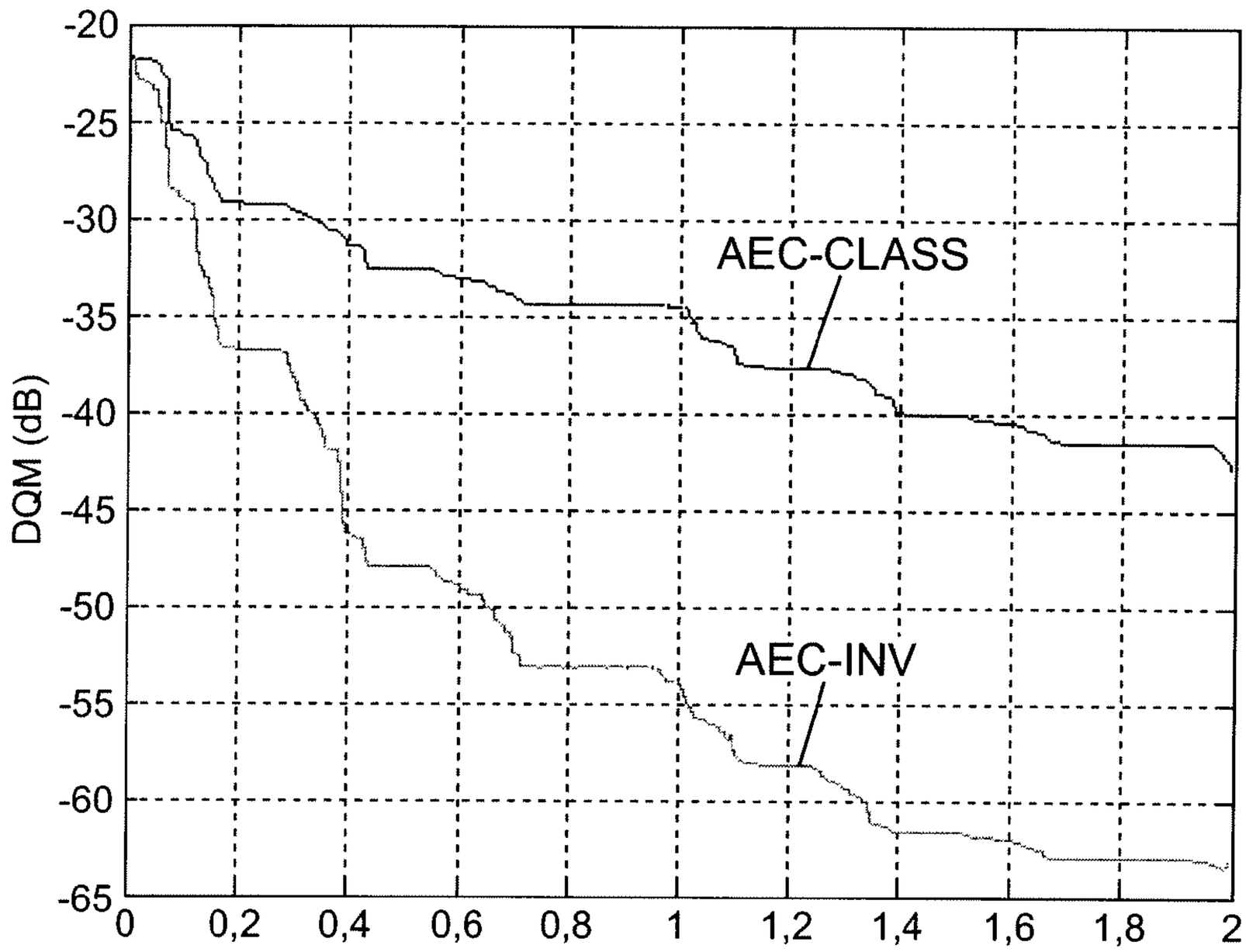

-20
-25
-30
-35
-40
-45
-50
-55
-60
-65
DQM (dB)
AEC-CLASS
AEC-INV
0
0,2
0,4
0,6
0,8
1
1,2
1,4
1,6
1,8
2
Ech
x10^5

FIG. 7

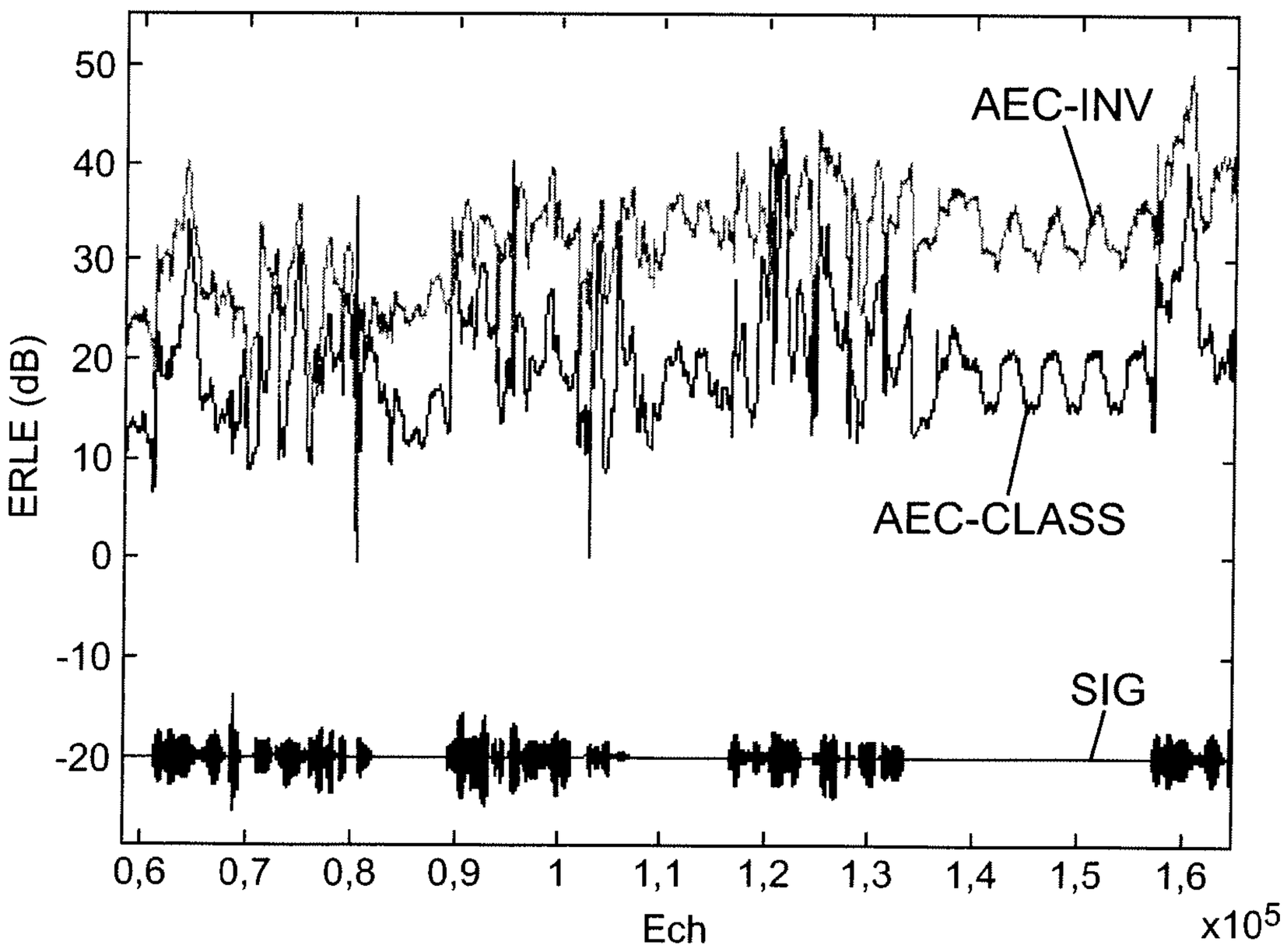

50
40
30
20
10
0
-10
-20
ERLE (dB)
AEC-INV
AEC-CLASS
SIG
0,6
0,7
0,8
0,9
1
1,1
1,2
1,3
1,4
1,5
1,6
Ech
x10^5

FIG. 8

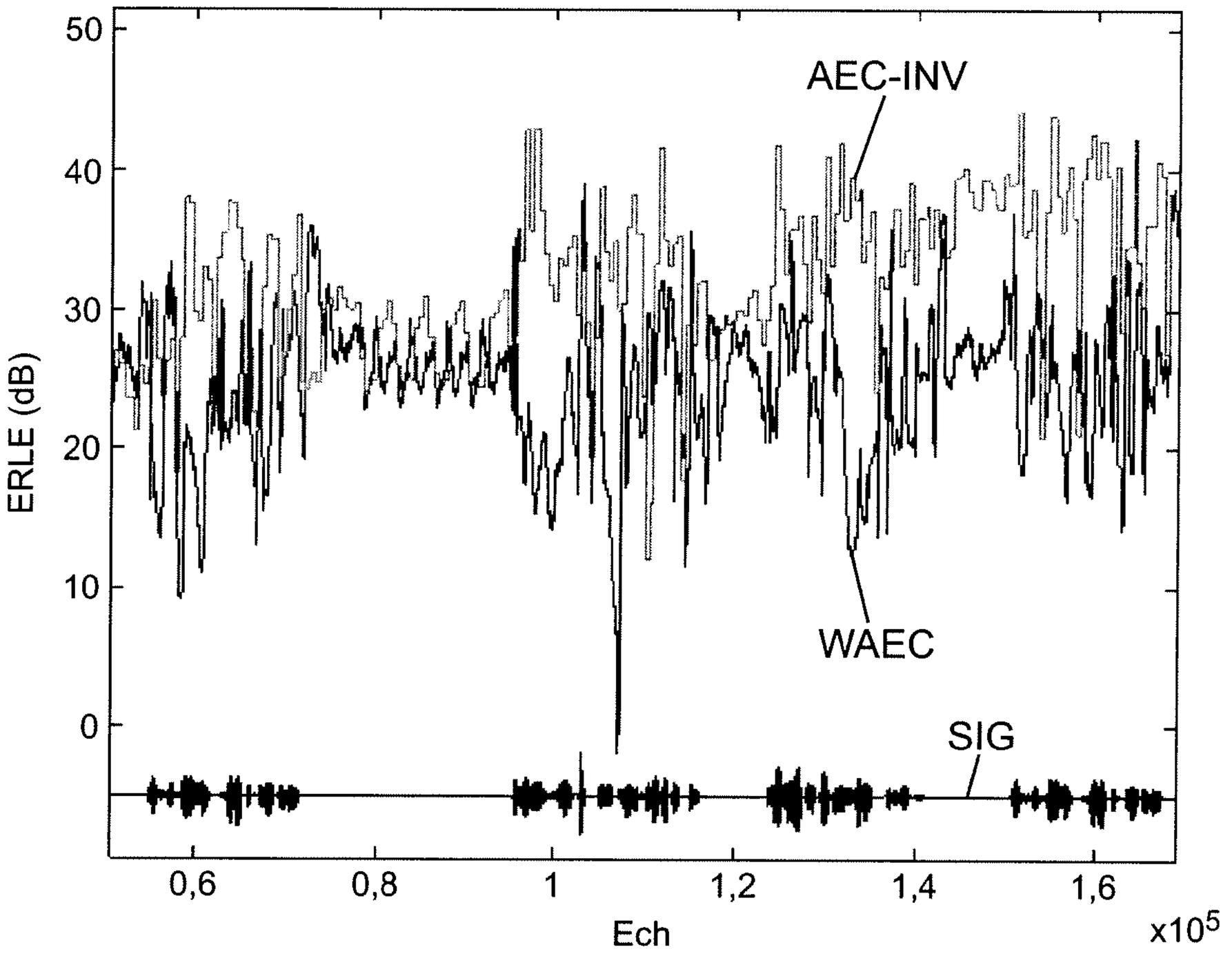
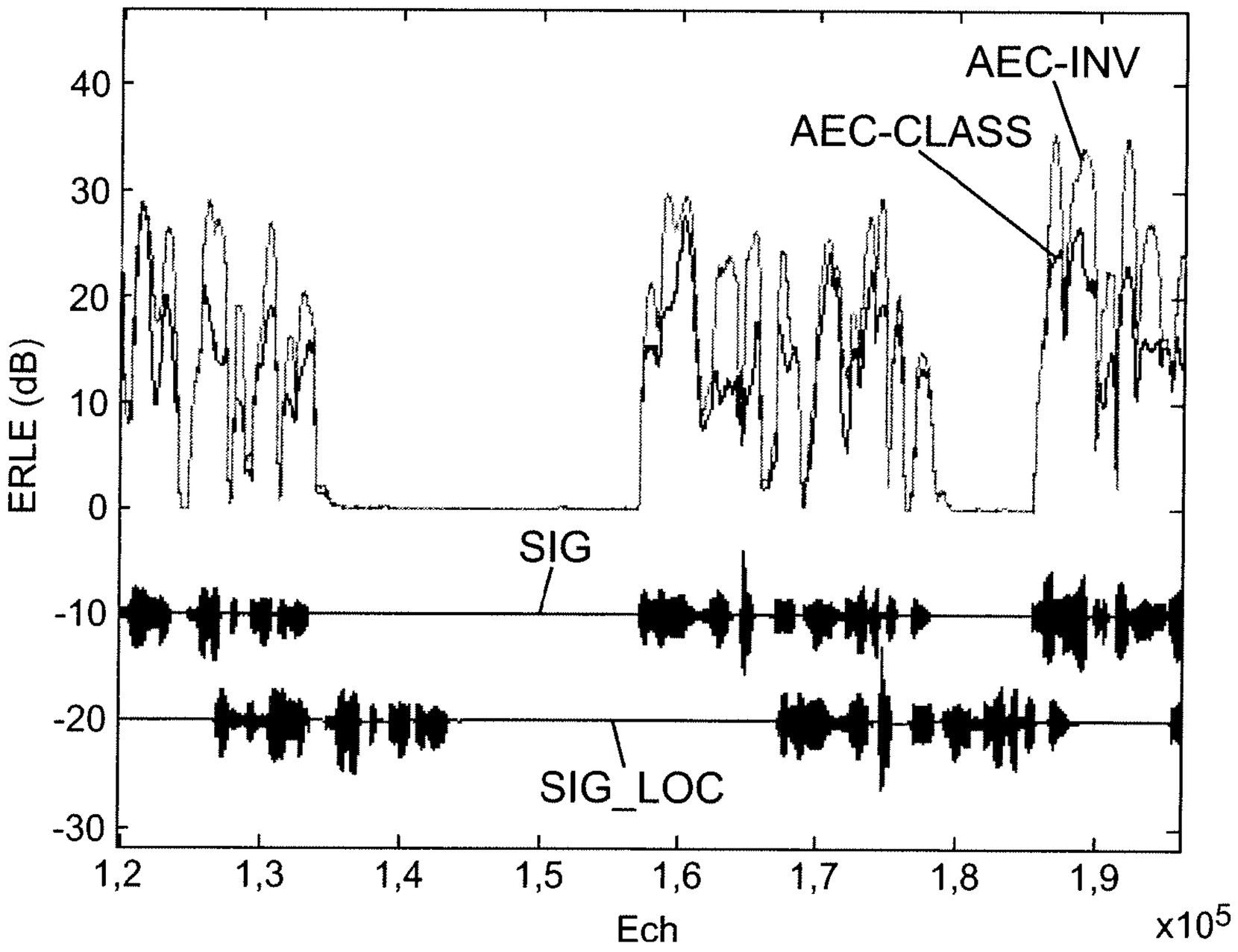

AEC-INV
WAEC
SIG
ERLE (dB)
50
40
30
20
10
0
0,6
0,8
1
1,2
1,4
1,6
Ech
x10^5

FIG. 9


AEC-INV
AEC-CLASS
SIG
SIG_LOC
ERLE (dB)
40
30
20
10
0
-10
-20
-30
1,2
1,3
1,4
1,5
1,6
1,7
1,8
1,9
Ech
x10^5

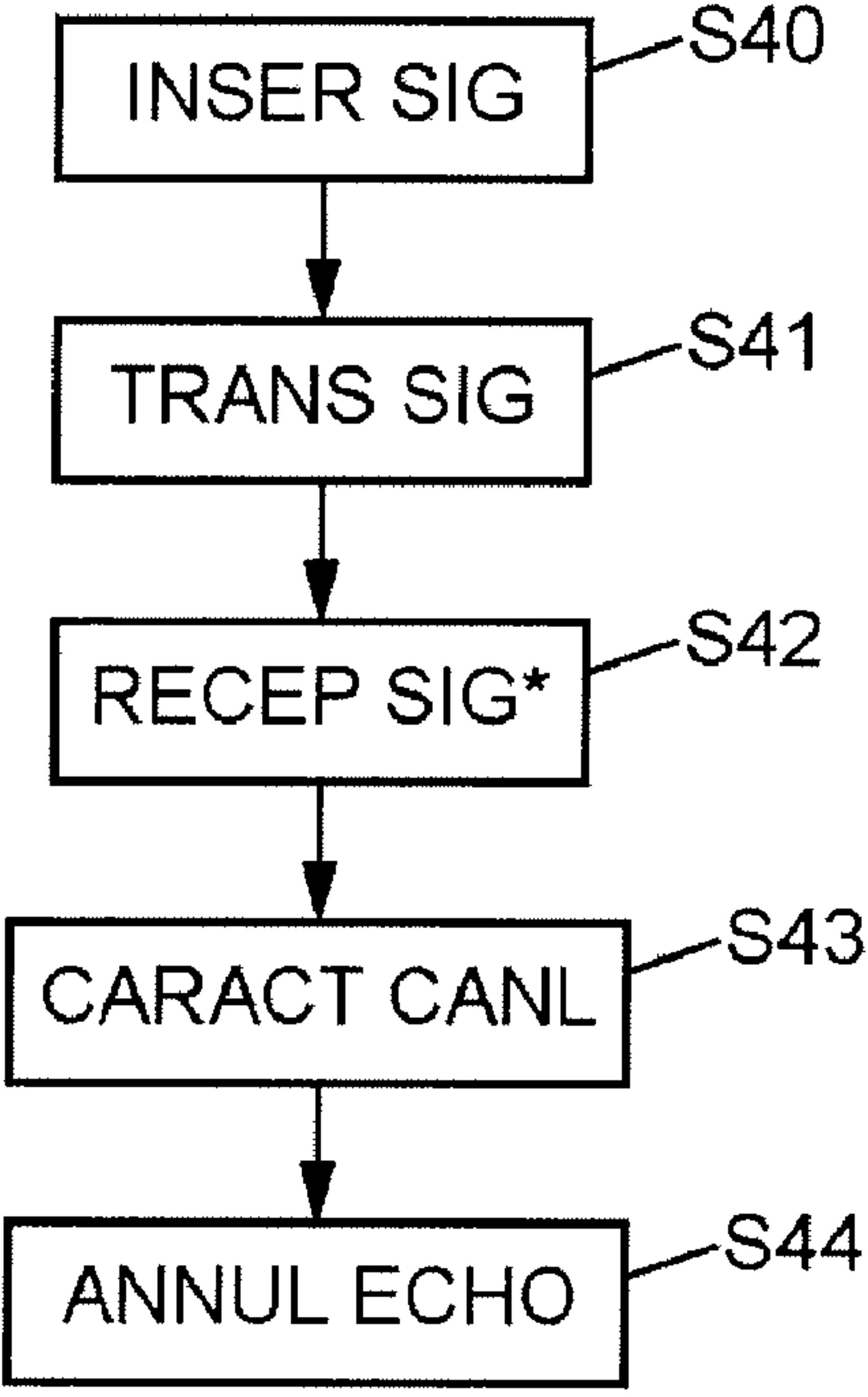

INSER SIG
S40
TRANS SIG
S41
RECEP SIG*
S42
CARACT CANL
S43
ANNUL ECHO
S44

FIG. 4