



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) **EP 1 143 416 A2**

(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:
10.10.2001 Patentblatt 2001/41

(51) Int Cl.7: **G10L 21/02**

(21) Anmeldenummer: **01440083.2**

(22) Anmeldetag: **22.03.2001**

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE TR**
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(72) Erfinder: **Walker, Michael**
73666 Baltmannsweiler 2 (DE)

(74) Vertreter: **Müller, Joachim, Dipl.-Ing. et al**
Alcatel
Intellectual Property Department, Stuttgart
70430 Stuttgart (DE)

(30) Priorität: **08.04.2000 DE 10017646**

(71) Anmelder: **ALCATEL**
75008 Paris (FR)

(54) **Geräuschunterdrückung im Zeitbereich**

(57) Ein Verfahren zur Geräuschreduktion bei der Übertragung von akustischen Nutzsignalen umfasst folgende Schritte:

- (a) Feststellen, wann eine Sprachpause vorliegt;
- (b) Abzweigen des ankommenden TK-Signals von Hauptsignalpfad und Anwenden einer Fourier-Transformation zur Erzeugung eines Frequenzspektrums;
- (c) Speichern des letzten während der letzten Sprachpause aufgenommenen Frequenzspektrums in einem Zwischenspeicher (3);
- (d) Anwenden einer inversen Fourier-Transformation auf das jeweils letzte aufgenommenen Frequenzspektrum zur Erzeugung eines nachgebilde-

ten Geräuschsignals;

(e) Abziehen des nachgebildeten Geräuschsignals im Zeitbereich vom aktuell ankommenden TK-Signal.

Dadurch bleibt das Originalsignal bis zum eigentlichen Geräuschabzug unverfälscht erhalten. Das Verfahren ermöglicht mit einfachen Mitteln und geringerem Rechenaufwand als bisher, einen für das menschliche Ohr möglichst angenehmen akustischen Gesamteindruck zu erzeugen, der an individuelle Bedürfnisse angepasst werden kann. Unabhängig von den Anforderungen für eine Sprachsignalverarbeitung kann eine einfache Optimierung auf die Erfordernisse der Spektralverarbeitung von Geräuschsignalen durchgeführt werden.

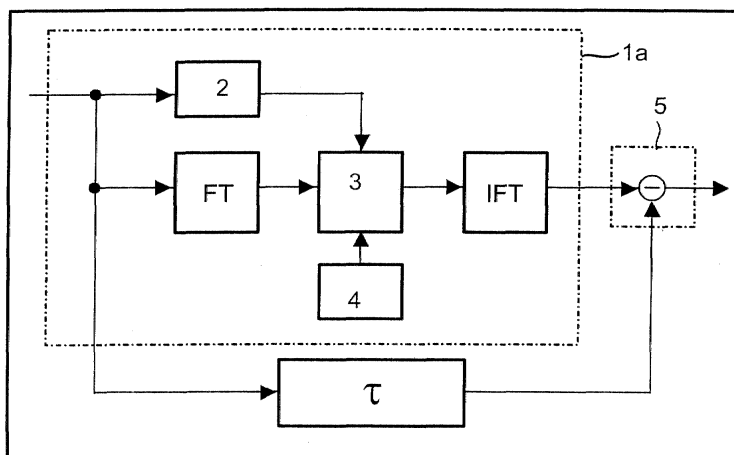


Fig.2

EP 1 143 416 A2

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren zur Reduktion von Geräuschsignalen bei Telekommunikations (=TK)-Systemen für die Übertragung von akustischen Nutzsignalen, insbesondere menschlicher Sprache.

[0002] Ein bekanntes Verfahren zur Geräuschreduktion ist die sogenannte "spektrale Subtraktion", die beispielsweise in der Veröffentlichung "A new approach to noise reduction based on auditory masking effects" von S. Gustafsson und P. Jax, ITG-Fachtagung, Dresden, 1998, beschrieben ist. Dabei handelt es sich um ein spektrales Geräuschabsenkungsverfahren, bei dem eine akustische Verdeckungsschwelle (beispielsweise nach dem MPEG-Standard) berücksichtigt wird.

[0003] Während einer natürlichen Kommunikation zwischen Menschen passt man in der Regel die Amplitude der gesprochenen Sprache automatisch an die akustische Umgebung an. Bei einer Sprachkommunikation zwischen entfernten Orten jedoch befinden sich die Gesprächspartner nicht im selben akustischen Umfeld und sind sich daher jeweils nicht der akustischen Situation am Ort des anderen Gesprächspartners bewusst. Verstärkt tritt daher ein Problem auf, wenn einer der Partner aufgrund seiner akustischen Umgebung gezwungen ist, sehr laut zu sprechen, während der andere Partner in einer leisen akustischen Umgebung Sprachsignale mit geringer Amplitude erzeugt.

[0004] Besonders verschärft ergeben sich Geräuschprobleme bei neueren Anwendungen von Kommunikationssystemen, wie beispielsweise Mobiltelefone, bei denen die Endgeräte derart klein gestaltet werden, dass eine unmittelbare räumliche Nachbarschaft zwischen Lautsprecher und Mikrofon nicht zu vermeiden ist. Aufgrund der direkten Schallübertragung, insbesondere durch Körperschall zwischen Lautsprecher und Mikrofon kann das akustische Störsignal in die gleiche Größenordnung kommen wie das Nutzsignal des Sprechers am jeweiligen Endgerät oder dieses sogar in der Amplitude übertreffen. Ein derartiges Geräuschproblem tritt auch bei mehreren räumlich benachbart angeordneten Endgeräten, beispielsweise in einem Büro- oder Konferenzraum mit vielen Telefonanschlüssen in nicht unerheblichem Umfang auf, da eine Kopplung von jedem Lautsprechersignal auf jedes Mikrofon erfolgt.

[0005] Hinzu kommt das Problem, dass auf einem TK-Kanal auch ein "elektronisch erzeugtes" Geräusch entsteht und als Hintergrund zum Nutzsignal mitübertragen wird. Um den Komfort beim Telefonieren zu erhöhen, ist man deshalb bestrebt, jede Art von Geräusch im Verhältnis zum Nutzsignal möglichst gering zu halten.

[0006] Schließlich ist man auch bestrebt, Störsignale wie unerwünschten Hintergrundlärm (Straßenlärm, Fabriklärm, Bürolärm, Kantinenlärm, Fluglärm etc.) zu reduzieren oder ganz zu unterdrücken.

[0007] Bei dem bekannten Kompander-Verfahren, wie es beispielsweise in der DE 42 29 912 A1 beschrieben

ben ist, wird der Grad der Geräuschabsenkung gemäß einer fest vorgegebenen Transferfunktion festgelegt. Der Kompander hat zunächst die Eigenschaft, Sprachsignale mit einem bestimmten (vorab eingestellten) "normalen Sprachsignalpegel" (ggf. normale Lautstärke genannt) praktisch unverändert von seinem Eingang zum Ausgang zu übertragen. Wird nun aber das Eingangssignal einmal zu laut, z.B. weil ein Sprecher zu dicht an sein Mikrofon kommt, so begrenzt ein Dynamik-Kompressor den Ausgangspegel auf nahezu den gleichen Wert wie im Normalfall, indem die aktuelle Verstärkung im Kompander mit zunehmender Eingangslautstärke linear abgesenkt wird. Durch diese Eigenschaft bleibt die Sprache am Ausgang des Kompandersystems etwa gleich laut - unabhängig davon wie stark die Eingangslautstärke schwankt. Wird nun andererseits ein Signal mit einem Pegel, der kleiner als der Normalpegel ist, auf den Eingang des Kompanders gegeben, so wird das Signal zusätzlich gedämpft, indem die Verstärkung zurückgeregelt wird, um Hintergrundgeräusche möglichst nur abgeschwächt zu übertragen. Der Kompander besteht somit aus zwei Teilfunktionen, einem Kompressor für Sprachsignalpegel, die größer oder gleich einem Normalpegel sind und einem Expander für Signalpegel, die kleiner als der Normalpegel sind.

[0008] Bei der oben angesprochenen spektralen Subtraktion wird zu diesem Zweck zunächst das Geräusch in den Sprachpausen gemessen und in Form eines Leistungsdichtespektrums fortlaufend in einem Speicher abgelegt. Das Leistungsdichtespektrum wird über eine Fourier Transformation gewonnen. Beim Auftreten von Sprache wird das gespeicherte Geräuschspektrum "als bester aktueller Schätzwert" vom aktuellen gestörten Sprachspektrum subtrahiert, sodann in den Zeitbereich zurücktransformiert, um auf diese Weise eine Geräuschreduktion für das gestörte Signal zu erhalten.

[0009] Nachteilig bei derartigen Verfahren ist die aufwendige Bestimmung dieser akustischen Verdeckungsschwelle und die Ausführung aller mit diesem Verfahren verbundenen Rechenoperationen. Ein weiterer Nachteil der spektralen Subtraktion besteht darin, dass durch den Vorgang einer prinzipiell nicht genauen spektralen Geräuschschätzung und nachfolgender Subtraktion auch Fehler im Ausgangssignal auftreten, die sich als "musical tones" bemerkbar machen.

[0010] Bei der erweiterten spektralen Signalbearbeitung, die ebenfalls in dem eingangs genannten Zitat beschrieben ist, werden zunächst mit Hilfe einer spektralen Subtraktion die Leistungsdichtespektren für das Geräusch und für die Sprache selbst geschätzt. Aus der Kenntnis dieser Teilspektren wird sodann mit Hilfe z.B. der Regeln aus dem MPEG Standard eine spektrale akustische Verdeckungsschwelle $R_T(f)$ für das menschliche Ohr berechnet. Mit Hilfe dieser Verdeckungsschwelle und den geschätzten Spektren für Geräusch und Sprache wird nach einer einfachen Regel eine Filterdurchlasskurve $H(f)$ berechnet, die so gestaltet ist,

dass wesentliche spektrale Teile der Sprache möglichst unverändert durchgelassen und spektrale Teile des Geräusches möglichst abgesenkt werden.

[0011] Sodann wird das originale gestörte Sprachsignal nur durch dieses Filter gegeben, um auf diese Weise eine Geräuschreduktion für das gestörte Signal zu erhalten. Der Vorteil dieses Verfahrens besteht nun darin, dass vom gestörten Signal "Nichts addiert oder subtrahiert" wird und daher Fehler in den Schätzungen weniger bis kaum wahrnehmbar sind. Nachteilig ist wieder der erheblich größere Rechenaufwand.

[0012] Von besonderem Nachteil bei allen diesen bekannten Verfahren ist die Tatsache, dass das eingehende Originalsignal bereits vor der eigentlichen Subtraktion eines wie auch immer nachgebildeten Geräuschsignals einem Signalbearbeitungsprozess unterworfen und damit prinzipiell verfälscht wird.

[0013] Aufgabe der vorliegenden Erfindung ist es demgegenüber, ein Verfahren möglichst geringer Komplexität mit den eingangs beschriebenen Merkmalen vorzustellen, bei dem in technisch unaufwendiger Art und Weise eine Geräuschabsenkung bzw. Geräuschunterdrückung erreicht wird, und bei der das Originalsignal bis zum eigentlichen Geräuschabzug unverfälscht erhalten bleibt. Dabei sollte das Verfahren mit einfachen Mitteln, insbesondere mit geringerem Rechenaufwand als bisher ermöglichen, einen für das menschliche Ohr möglichst angenehmen akustischen Gesamteindruck zu erzeugen, der je nach Geschmack an individuelle Bedürfnisse angepasst werden kann. Schließlich sollte das neue Verfahren völlig unabhängig von den Anforderungen für eine Sprachsignalverarbeitung durchgeführt werden können und damit eine einfache Optimierung auf die Erfordernisse der Spektralverarbeitung von Geräuschsignalen ermöglichen.

[0014] Erfindungsgemäß wird diese Aufgabe auf ebenso einfache wie wirkungsvolle Art und Weise durch folgende Verfahrensschritte gelöst:

(a) Feststellen mittels Sprach-Pausen-Detektion, wann in der zu übertragenden Mischung aus Nutzsensignalen und Störsensignalen ein Sprachsignal enthalten ist oder wann eine Sprachpause vorliegt;

(b) Abzweigen des ankommenden TK-Signals von Hauptsignalpfad und Anwenden einer Fourier-Transformation auf das abgezweigte TK-Signal zur Erzeugung eines Frequenzspektrums des abgezweigten TK-Signals;

(c) Speichern des letzten während der letzten Sprachpause aufgenommenen Frequenzspektrums in einem Zwischenspeicher;

(d) Anwenden einer inversen Fourier-Transformation auf das jeweils letzte aufgenommenen Frequenzspektrum zur Erzeugung eines nachgebildeten Geräuschsignals;

(e) Abziehen des nachgebildeten Geräuschsignals im Zeitbereich vom aktuell ankommenden TK-Signal.

[0015] Das erfindungsgemäße Verfahren ermöglicht durch die getrennte Nachbildung des Geräuschsignals im Frequenzbereich unabhängig von einer Verarbeitung des originalen Sprachsignals einen direkten Abzug des nachgebildeten Geräuschsignals vom originalen, unverfälschten Eingangssignal, welches weder einer Fourier-Transformation noch einer inversen Fourier-Transformation unterworfen wird. Bei einer entsprechenden Phasenkorrektur im Frequenzbereich ist sogar eine Geräuschsubtraktion vom Originalsignal ohne Zeitverzögerung möglich. Dabei ist das erfindungsgemäße Verfahren weniger komplex als die oben geschilderten bekannten Verfahren aus dem Stand der Technik, erfordert geringere Rechenleistungen und führt zu einer besseren Frequenzauflösung.

[0016] Durch die Abtrennung der Geräuschnachbildung von der Weitergabe des Originalsignals ermöglicht das erfindungsgemäße Verfahren in einer besonders bevorzugten Variante, dass in Schritt (d) nur ein ausgewählter Teil des erzeugten Frequenzspektrums zur Erzeugung des nachgebildeten Geräuschsignals verwendet wird. Damit kann die zur Durchführung des erfindungsgemäßen Verfahrens erforderliche Rechenleistung weiter minimiert bzw. das Verfahren selbst noch schneller durchgeführt werden.

[0017] Eine Weiterbildung dieser Verfahrensvariante zeichnet sich dadurch aus, dass die Auswahl des zur Erzeugung des nachgebildeten Geräuschsignals verwendeten Teils des Frequenzspektrums nach Kriterien der Psychoakustik gemäß den Mittelwerten des Wahrnehmungsspektrums des menschlichen Gehörs erfolgt.

[0018] Dabei wird der Wert für das nachzubildende Geräuschsignal nicht nur aus dem momentanen Leistungswert eines Originalsignals in Sprachpausen alleine, sondern auch aus einem gewichteten spektralen Verlauf des entsprechenden Signals bestimmt und insgesamt über die so gewonnene Funktion eine gehörliche, d.h. eine psychoakustisch angenehm klingende Geräuschreduktion erzielt.

[0019] Da es kein einfach darstellbares Maß für eine akustisch angenehm klingende Geräuschreduktion gibt, sind alle Qualitätsbeurteilungen auf umfangreiche Gehörtests angewiesen, die anschließend mittels dafür optimierter statistischer Methoden ausgewertet werden, um einen Bewertungsmaßstab, (ähnlich wie bei Sprachcodern) zu erhalten.

[0020] Die grundlegenden Vorgehensweisen hierzu sind beispielsweise dem Lehrbuch von E. Zwicker, "Psychoakustik", Springer-Verlag Berlin, 1982, insbesondere Seiten 51-53 zu entnehmen.

[0021] Durch die psychoakustische Bewertung kann nicht nur die wahrnehmbare Qualität des Gesamtsignals optimiert werden, sondern es sind zudem auch

noch weitere Einsparungen in der erforderlichen Rechenleistung möglich, wenn beispielsweise Maskiereffekte ausgenutzt werden oder nur diejenigen Frequenzen Berücksichtigung finden, die eindeutig durch Geräusch- oder Störquellen verursacht wurden.

[0022] In einer alternativen Weiterbildung der obigen Verfahrensvariante erfolgt die Auswahl des zur Erzeugung des nachgebildeten Geräuschsignals verwendeten Teiles des Frequenzspektrums derart, dass nur diskrete Frequenzen des Spektrums betrachtet werden, und dass der Abstand der diskreten Frequenzen in Richtung höherer Frequenzen stetig größer, vorzugsweise nach einer logarithmischen Funktion gewählt wird. Damit ist die Frequenzauflösung an die Wahrnehmung des menschlichen Ohres besser angepasst.

[0023] Weiter verbessert werden können diese Weiterbildungen dadurch, dass der ausgewählte Teil des Frequenzspektrums in vorher festgelegte Frequenzgruppen aufgeteilt wird, und dass in jeder Frequenzgruppe nur die Frequenz bzw. das Frequenzband mit der größten Signalenergie innerhalb der Frequenzgruppe selektiert und zur Erzeugung des nachgebildeten Geräuschsignals weiterverwendet wird. Mit dieser Selektion wird eine starke Reduzierung der zu berechnenden Frequenzen bei gleichbleibender hörbarer oder wahrnehmbarer Qualität erreicht, wodurch die Rechenleistung für das Verfahren weiter verringert und die Qualität des Ausgangssignals noch weiter erhöht wird.

[0024] Besonders vorteilhaft ist es, wenn die Selektion der Frequenz bzw. des Frequenzbandes mit der größten Signalenergie innerhalb der Frequenzgruppe vor Schritt (c) bzw. vor Schritt (d) erfolgt. Durch die Auswahl einer bestimmten Frequenz aus einer Frequenzgruppe sind Unterschiede in der Signalenergie besonders einfach detektierbar.

[0025] Vorteilhaft ist auch eine Verfahrensvariante, bei der in Schritt (b) das Frequenzspektrum des abgezweigten TK-Signals nur in einem vorgegebenen Frequenzbereich erzeugt wird. Sofern die Störquelle nur ein eingeschränktes Frequenzspektrum aufweist, kann mit dieser Maßnahme wiederum erhebliche Rechenleistung eingespart werden. Beispielsweise ist in Kraftfahrzeugen mit Störquellen in einem Frequenzbereich nur bis zu maximal 1 KHz zu rechnen, da das Störsignal in der Hauptsache durch niederfrequente Schallerzeugung (Motor, Getriebe, Rollgeräusche etc.) gebildet wird.

[0026] Besonders einfach ist eine Verfahrensvariante, die sich dadurch auszeichnet, dass in Schritt (b) und/oder in Schritt (d) eine diskrete Fourier-Transformation bzw. eine inverse diskrete Fourier-Transformation angewandt wird, wobei vom ankommenden TK-Signal zeitdiskrete Amplitudenwerte mit einer Abtastfrequenz f_T abgetastet werden.

[0027] In einer bevorzugten Weiterbildung der Verfahrensvariante wird in Schritt (b) eine schnelle Fourier-Transformation (= FFT) angewandt. Wenn ein großes Frequenzgebiet bei gleichzeitiger hoher Frequenzauflö-

sung abgedeckt werden soll, lässt sich mit dieser Vorgehensweise die Analyse mit der geringsten Rechenleistung ausführen. Die FFT ist besonders dann sinnvoll, wenn beispielsweise mit mehr als 128 Frequenzlinien gerechnet werden muss.

[0028] Vorteilhafterweise kann in Schritt (d) eine inverse diskrete Fourier-Transformation (= IDFT) angewandt werden. Hierdurch lässt sich eine Signalsynthese mit der geringsten Rechenleistung ausführen, wenn ein ausgewähltes Spektrum bearbeitet wird, da der Nachteil einer äquidistanten Frequenzaufteilung bei der FFT vermieden wird. Die IDFT kann daher vorteilhaft für einen definierten Frequenzbereich angewendet werden. Die Verteilung der Frequenzen kann individuell erfolgen. Ab einer Frequenzauflösung von weniger als 128 Frequenzlinien ist eine Einsparung der Rechenleistung gegenüber der FFT möglich.

[0029] In der Applikation sind auch Einsparungen in der Rechenleistung oder Qualitätsverbesserungen erreichbar, wenn in Schritt (d) eine inverse schnelle Fourier-Transformation (=IFFT) angewandt wird. In Kombination mit einer FFT in Schritt (b) können breitbandige Störer besonders wirtschaftlich bearbeitet werden.

[0030] Alternativ zur letztgenannten Verfahrensvariante ist eine Ausführungsform, bei der nur der Teil des erzeugten Frequenzspektrums ausgewählt wird, der unterhalb der halben Abtastfrequenz $f_T/2$ liegt. Damit lassen sich wiederum Einsparungen in der Rechenleistung, aber auch im Speicherplatzaufwand erzielen.

[0031] Besonders vorteilhaft ist auch eine Variante des erfindungsgemäßen Verfahrens, bei der in Schritt (c) ein Frequenzspektrum zwischengespeichert wird, das durch eine Mittelung des aktuell in Schritt (b) erzeugten Frequenzspektrums mit vorher erzeugten Frequenzspektren gewonnen wird. Durch die Mittelung werden Spektrallinien mit großer Energie gefunden und Zufallswerte oder sporadische Fehler systematisch unterdrückt.

[0032] Dabei ist es besonders günstig, wenn die Mittelung mit unterschiedlicher relativer Wichtung des aktuell erzeugten Frequenzspektrums in unterschiedlichen Frequenzbereichen erfolgt. Allgemein kann mit derartigen unterschiedlichen Richtungen das natürliche Einschwingverhalten von Störern berücksichtigt werden. Beispielsweise kann sich die Drehzahl eines Motors in einem Kraftfahrzeug in der Regel nicht schlagartig ändern. Niederfrequente Störer haben eine höhere Einschwingzeit als hochfrequente. Die vorgeschlagene Wichtung hilft dabei, die Adaptivität eines Systems stabil und schnell zu machen.

[0033] Dabei ist es wiederum besonders vorteilhaft, wenn die Wichtung nach Kriterien der Psychoakustik gemäß den Mittelwerten des Wahrnehmungsspektrums des menschlichen Gehörs erfolgt. Wie bereits oben diskutiert, werden bei der psychoakustischen Wichtung die frequenzabhängigen Einschwingzeiten dem menschlichen Hörempfinden angepasst. Dadurch erreicht man eine Optimierung des Systems hinsichtlich Natürlich-

keit, Stabilität und Adaptionszeit.

[0034] Um eine Überkompensation bei der Geräuschbehandlung zu vermeiden, wird bei einer besonders bevorzugten Variante des erfindungsgemäßen Verfahrens in Schritt (e) ein nach vorgegebenen Kriterien mit einem Wichtungsfaktor $a < 1$ gewichtetes nachgebildetes Geräuschsignal vom aktuell ankommenden TK-Signal abgezogen.

[0035] Bei einer vorteilhaften Weiterbildung wird der Wichtungsfaktor a als ein von Fehlern des TK-Systems abhängiger konstanter Wert gewählt. Dies ermöglicht eine kostengünstige und einfache Optimierung des erfindungsgemäßen Verfahrens an die Fehler des jeweiligen TK-Systems. Werden die Fehler automatisch erfasst, so kann die Wichtung auch während des Betriebs erfolgen.

[0036] Alternativ kann der Wichtungsfaktor a als ein nach einem vom Benutzer des TK-Systems wählbaren Qualitätsmaß einstellbarer Wert gewählt werden. Ein derartiger vom Benutzer definierter Wichtungsfaktor ermöglicht eine individuelle, benutzerdefinierte Anpassung des erfindungsgemäßen Verfahrens an die individuellen Bedürfnisse. Wird das erfindungsgemäße System in ein vorhandenes übergeordnetes Konzept integriert, kann ein vom Benutzer bereitgestellter statistischer Wert, wie beispielsweise die Fehlerwahrscheinlichkeit oder Erkennungsrate zur Steuerung des Wichtungsfaktors herangezogen werden. Bei Anwendungen im Kraftfahrzeugbereich kann der Wichtungsfaktor beispielsweise auch aus der Drehzahl oder Geschwindigkeit abgeleitet werden.

[0037] Dies lässt sich dadurch weiterverbessern, dass der Wichtungsfaktor a adaptiv an das aktuelle ankommende TK-Signal angepasst wird. Die adaptive Gewichtung erlaubt eine automatische Optimierung der Geräuschminderung während des Betriebes. Der Wichtungsfaktor kann von statistischen Werten wie Fehlerwahrscheinlichkeit, Mittelwert, Zustandsänderungen etc. abgeleitet werden. Mit der adaptiven Wichtung sind besonders einfache und schnelle Anpassungen des erfindungsgemäßen Verfahrens an individuelle Gegebenheiten in der akustischen Umgebung des TK-Endgeräts möglich.

[0038] Eine weitere vorteilhafte Variante des erfindungsgemäßen Verfahrens zeichnet sich dadurch aus, dass dem in Schritt (d) erzeugten nachgebildeten Geräuschsignal vor Schritt (e) ein synthetisches Geräuschsignal beigemischt wird. Die Beimischung eines künstlichen Rauschsignals mit konstanter Leistungsdichte kann zur Maskierung von dynamischen, nicht-stationären Störern im Ausgangssignal dienen.

[0039] Eine weitere Variante des erfindungsgemäßen Verfahrens sieht vor, dass das aktuell ankommende TK-Signal vor Schritt (e) einer definierten Zeitverzögerung unterworfen wird, die vorzugsweise so ausgelegt ist, dass die Phasenlage des ankommenden TK-Signals mit der Phasenlage des nachgebildeten Geräuschsignals vor dem Abzug übereinstimmt.

[0040] Bei einer alternativen Verfahrensvariante ist vorgesehen, dass das aktuell ankommende TK-Signal unverzüglich dem Abzug in Schritt (e) zugeführt wird, und dass das nachgebildete Geräuschsignal in seiner Phasenlage vor Schritt (e) an die Phasenlage des aktuell ankommenden TK-Signals angepasst wird. Wird die Phasenlage des reproduzierten Geräuschsignals im Frequenzbereich vor der Rücktransformation korrigiert, kann die Subtraktion vom unverzögerten Signal im Zeitbereich erfolgen. Störende Signalverzögerungen können somit entfallen. Diese treten zwangsläufig bei allen Verfahren auf, bei denen das Nutzsignal (Sprache) den Umweg über zwei Transformationen macht, wie beispielsweise bei der oben diskutierten bekannten spektralen Subtraktion.

[0041] Besonders bevorzugt ist eine Variante des erfindungsgemäßen Verfahrens, bei der zusätzlich zur Erkennung und Reduktion von Geräuschsignalen das Vorliegen von Echosignalen detektiert und/oder vorhergesagt wird und die Echosignale unterdrückt oder reduziert werden. Eine zusätzliche Echounterdrückung ist allerdings nur dann möglich, wenn das empfangene Originalsignal vom fernen TK-Teilnehmer zur Echoberechnung miteinbezogen wird. Dies bedeutet, dass die Geräuschreproduktion auch eine Echoreproduktion beinhaltet, die mit einem vom fernen TK-Teilnehmer ankommenden Signal verbunden ist.

[0042] Diese Verfahrensvariante kann dadurch verbessert werden, dass die Steuerung der Reduktion von Geräuschsignalen und der Reduktion von Echosignalen getrennt erfolgt.

[0043] Vorteilhaft ist es auch, wenn während der Zeitdauer einer Echo-Reduktion zum Nutzsignal zusätzlich ein künstliches Geräuschsignal addiert wird, wie es bereits oben näher diskutiert wurde, um den subjektiven Eindruck einer "toten Leitung" zu vermeiden.

[0044] Insbesondere kann das künstliche Geräuschsignal eine psychoakustisch als angenehm empfundene akustische Signalsequenz (=comfort noise) umfassen.

[0045] Alternativ kann das künstliche Geräuschsignal ein zuvor während der aktuellen TK-Verbindung aufgezeichnetes Geräuschsignal umfassen, das die aktuelle akustische Umgebungssituation besonders "lebensecht" nachzubilden vermag.

[0046] In den Rahmen der vorliegenden Erfindung fällt auch eine Servereinheit, eine Prozessor-Baugruppe sowie eine Gate-Array-Baugruppe zur Unterstützung des oben beschriebenen erfindungsgemäßen Verfahrens sowie ein Computerprogramm zur Durchführung des Verfahrens. Das Verfahren kann sowohl als Hardwareschaltung, als auch in Form eines Computerprogramms realisiert werden. Heutzutage wird eine Software-Programmierung für leistungsstarke DSP's bevorzugt, da neue Erkenntnisse und Zusatzfunktionen leichter durch eine Veränderung der Software auf bestehender Hardwarebasis implementierbar sind. Verfahren können aber auch als Hardwarebausteine beispielsweise

se in TK-Endgeräten oder Telefonanlagen implementiert werden.

[0047] Weitere Vorteile der Erfindung ergeben sich aus der Beschreibung und der Zeichnung. Ebenso können die vorstehend genannten und die noch weiter aufgeführten Merkmale erfindungsgemäß jeweils einzeln für sich oder zu mehreren in beliebigen Kombinationen Verwendung finden. Die gezeigten und beschriebenen Ausführungsformen sind nicht als abschließende Aufzählung zu verstehen, sondern haben vielmehr beispielhaften Charakter für die Schilderung der Erfindung.

[0048] Die Erfindung ist in der Zeichnung dargestellt und wird anhand von Ausführungsbeispielen näher erläutert. Es zeigen:

Fig. 1 ein stark schematisiertes Diagramm der Funktionsweise einer Einrichtung zur Durchführung des erfindungsgemäßen Verfahrens;

Fig. 2 eine detailliertere schematische Darstellung einer Einrichtung zur Durchführung des erfindungsgemäßen Verfahrens;

Fig. 3 ein Schema für ein spektrales Subtraktionsverfahren nach dem Stand der Technik;

Fig. 4 eine Ausführungsform der Erfindung mit schneller Fourier-Transformation und schneller Rücktransformation sowie blockweise überlappender Bearbeitung des eingegebenen Zeitsignals im Frequenzbereich;

Fig. 5 Ein Schema einer Ausführungsform mit gleichzeitiger Echo-reduktion;

Fig. 6a ein Beispiel eines mit FFT berechneten Geräuschsignals im Frequenzraum;

Fig. 6b ein mit einer diskreten Fourier-Transformation und nur bis $f_s/2$ berechneten Geräuschsignals; und

Fig. 6c ein Geräuschsignal im Frequenzbereich bis $f_s/2$ als Ergebnis einer modifizierten Fourier-Transformation mit höherer Auflösung.

[0049] In Fig. 1 ist gezeigt, wie aus einem ankommenden Originalsignal x , welches einen Sprachanteil s sowie einen Geräuschanteil n enthält, einerseits in einer Einrichtung 1 ein Geräuschsignal y_n im Frequenzbereich nachgebildet wird und andererseits das Originalsignal x_{s+n} getrennt von der Geräuschnachbildung einer Geräuschsubtraktion zugeführt wird, wobei wahlweise eine Zeitverzögerung τ vorgenommen werden kann. Das geräuschreduzierte Signal y_s wird dann im TK-System weitergeleitet.

[0050] In Fig. 2 ist eine einfache Ausführungsform ge-

zeigt, bei der in der Einrichtung 1a zur Geräuschnachbildung ein praktisch immer erforderlicher Sprachpausendetektor 2 vorgesehen ist, mit dem ermittelt wird, wann das eingehende Signal Sprachsignale enthalten kann oder wann eine Sprachpause vorliegt. Parallel dazu wird das eingehende TK-Signal einer Fourier-Transformation FT zur Erzeugung eines Frequenzspektrums unterworfen und das jeweils daraus entstehende Frequenzspektrum in einem Zwischenspeicher 3 abgespeichert. Die zeitlich nacheinander abgespeicherten Frequenzspektren können mit Hilfe einer Einrichtung 4 einer Mittelwertbildung unterzogen werden.

[0051] Sobald der Sprachpausendetektor 2 feststellt, dass eine Sprachpause beendet ist und im eingehenden Originalsignal auch Sprachsignale vorhanden sein können, wird das letzte im Zwischenspeicher 3 abgespeicherte Frequenzspektrum (ggf. gemittelt mit vorher aufgenommenen Spektren) einer inversen Fourier-Transformation IFT unterworfen und in einem Subtraktionsglied 5 vom Originalsignal, das ggf. einer Zeitverzögerung τ unterworfen wurde, abgezogen, um ein geräuschbefreites oder zumindest geräuschreduziertes Signal zu erhalten.

[0052] Im Gegensatz dazu wird bei bekannten Verfahren der spektralen Subtraktion das eingehende Originalsignal, wie in Fig. 3 dargestellt, direkt einer Fourier-Transformation FT unterzogen, ein nachgebildetes Geräuschsignal im Frequenzbereich in einem Subtraktionsglied 5' vom Fourier-transformierten Originalsignal abgezogen und das daraus entstehende neue, geräuschreduzierte Signal im Frequenzbereich einer inversen Fourier-Transformation

[0053] IFT unterzogen und als geräuschreduziertes TK-Signal im Zeitbereich weitergeleitet. Es findet also bei den bekannten Verfahren im Stand der Technik grundsätzlich immer eine Veränderung des Originalsignals auch schon vor dem eigentlichen Geräuschabzug statt.

[0054] In Fig. 4 ist eine weitere Ausführungsform der Erfindung dargestellt, bei der das zunächst im Zeitbereich eingehende Originalsignal x_{s+n} in der Einrichtung 1b zur Geräuschnachbildung blockweise verarbeitet wird. Hierbei wird das Zeitsignal vor der Transformation in den Frequenzbereich einer Fensterung (z.B. nach Hamming) in einer entsprechend vorgeschalteten Vorrichtung 4' bzw. 4'' unterzogen. Um bei der Rücktransformation die durch die Fensterung entstandenen Fehler zu kompensieren, wird neben der Verarbeitung in einem ersten Pfad eine parallele Verarbeitung in einem weiteren Pfad mit der gleichen Fensterung vorgenommen, wobei lediglich das Signal um die halbe Fensterlänge versetzt ist und ansonsten das nachzubildende Geräuschsignal mit den gleichen Mitteln berechnet wird, wodurch eine Kompensation der durch die Fensterung erzeugten Fehler erreicht werden kann.

[0055] Im Einzelnen wird bei dem gezeigten Beispiel im ersten Pfad die Fensterung in einer Vorrichtung 4' vorgenommen, danach das Zeitsignal einer schneller

Fourier-Transformation FFT unterworfen und das entstehende Spektrum in einem Zwischenspeicher 3' gespeichert. Das Gleiche geschieht im zweiten Pfad über eine Fenstervorrichtung 4" und eine Zwischenspeicherung des Fourier-transformierten Signals in einem Zwischenspeicher 3". An die Zwischenspeicher 3', 3" schließt sich eine inverse schnelle Fourier-Transformation IFFT jeweils an, und die daraus hervorgehenden Spektren im Zeitbereich werden in einer Überlappeneinrichtung 6 zu einem nachgebildeten Geräuschsignal Y_n zusammengeführt. Anschließend wird das nachgebildete Geräuschsignal wiederum im Subtraktionsglied 5 von einem wahlweise um eine Zeit τ zeitversetzten Originalsignal X_{s+n} abgezogen, um das geräuschbereinigte Ausgangssignal y_s zu erhalten. Die Subtraktion des Geräuschsignals vom Originalsignal im Subtraktionsglied 5 kann phasenangepasst erfolgen.

[0056] Ein weiteres Ausführungsbeispiel ist in Fig. 5 dargestellt, wo das abgezweigte eingehende TK-Signal x_{s+n+e} neben Sprach- und Geräuschsignalen auch Echosignale enthält. In einer Einrichtung 1c zur Geräusch- und Echonachbildung wird außerdem ein Echosignal e eingegeben, welches in einem zum Geräuschnachbildungspfad parallelen Verarbeitungspfad weiterbehandelt wird.

[0057] Das eingehende Originalsignal x_{s+n+e} wird zunächst einer Fensterung in einer Vorrichtung 4a unterzogen, danach einer schnellen Fourier-Transformation FFT und das erhaltene Frequenzspektrum in einem Zwischenspeicher 3a zwischengespeichert. Parallel dazu wird das Echosignal e in einer Vorrichtung 4b ebenfalls einer Fensterung unterzogen und danach Fourier-transformiert. Die Frequenzspektren beider Pfade werden in einen Zwischenspeicher 3b zwischengespeichert und evtl. einer Mittelung unterzogen. Danach erfolgt auf beiden Pfaden jeweils getrennt eine schnelle inverse Fourier-Transformation IFFT. In einer Einrichtung 6a schließlich werden das nachgebildete Geräuschsignal und das nachgebildete Echosignal zu einem abziehenden Gesamtsignal y_{n+e} überlappt, welches in der Subtraktionseinrichtung 5 von dem unveränderten oder um eine Zeit τ verzögerten Originalsignal x_{s+n+e} abgezogen wird, um das geräusch- und echoreduzierte TK-Signal y_s zu erhalten.

[0058] Die Fign. 6a bis 6c schließlich zeigen Beispiele für nach dem erfindungsgemäßen Verfahren berechnete Geräuschsignale im Frequenzraum. Dabei ist bei dem Beispiel nach Fig. 6a das nachzubildende Geräuschsignal aus einer schnellen Fourier-Transformation FFT gewonnen worden. Man sieht die typische Spiegelsymmetrie um den halben Frequenzwert $f_s/2$.

[0059] Es reicht aber auch schon aus, wenn nur die erste Hälfte des nachgebildeten Geräuschsignals im Frequenzraum bis zur Frequenz $f_s/2$ verwendet wird, was in Fig. 6b anhand eines Beispiels dargestellt ist, dessen Ergebnis mit Hilfe einer diskreten Fourier-Transformation erhalten wurde.

[0060] Fig. 6c schließlich zeigt das Ergebnis der An-

wendung einer modifizierten diskreten Fourier-Transformation mit höherer Auflösung, wobei wiederum nur die Hälfte des Frequenzspektrums bis zur Frequenz $f_s/2$ verarbeitet wird.

Patentansprüche

1. Verfahren zur Reduktion von Geräuschsignalen bei Telekommunikations(=TK)-Systemen für die Übertragung von akustischen Nutzsignalen, insbesondere menschlicher Sprache, mit folgenden Schritten:

(a) Feststellen mittels Sprach-Pausen-Detektion, wann in der zu übertragenden Mischung aus Nutzsignalen und Störsignalen ein Sprachsignal enthalten ist oder wann eine Sprachpause vorliegt;

(b) Abzweigen des ankommenden TK-Signals von Hauptsignalpfad und Anwenden einer Fourier-Transformation auf das abgezweigte TK-Signal zur Erzeugung eines Frequenzspektrums des abgezweigten TK-Signals;

(c) Speichern des letzten während der letzten Sprachpause aufgenommenen Frequenzspektrums in einem Zwischenspeicher (3);

(d) Anwenden einer inversen Fourier-Transformation auf das jeweils letzte aufgenommenen Frequenzspektrum zur Erzeugung eines nachgebildeten Geräuschsignals;

(e) Abziehen des nachgebildeten Geräuschsignals im Zeitbereich vom aktuell ankommenden TK-Signal.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** in Schritt (d) nur ein ausgewählter Teil des erzeugten Frequenzspektrums zur Erzeugung des nachgebildeten Geräuschsignals verwendet wird.

3. Verfahren nach Anspruch 2, **dadurch gekennzeichnet, dass** die Auswahl des zur Erzeugung des nachgebildeten Geräuschsignals verwendeten Teils des Frequenzspektrums nach Kriterien der Psychoakustik gemäß den Mittelwerten des Wahrnehmungsspektrums des menschlichen Gehörs erfolgt.

4. Verfahren nach Anspruch 2, **dadurch gekennzeichnet, dass** die Auswahl des zur Erzeugung des nachgebildeten Geräuschsignals verwendeten Teiles des Frequenzspektrums derart erfolgt, **dass** nur diskrete Frequenzen des Spektrums betrachtet

werden, und **dass** der Abstand der diskreten Frequenzen in Richtung höherer Frequenzen stetig größer, vorzugsweise nach einer logarithmischen Funktion gewählt wird.

5. Verfahren nach Anspruch 2, **dadurch gekennzeichnet, dass** der ausgewählte Teil des Frequenzspektrums in vorher festgelegte Frequenzgruppen aufgeteilt wird, und **dass** in jeder Frequenzgruppe nur die Frequenz bzw. das Frequenzband mit der größten Signalenergie innerhalb der Frequenzgruppe selektiert und zur Erzeugung des nachgebildeten Geräuschsignals weiterverwendet wird. 5
6. Verfahren nach Anspruch 5, **dadurch gekennzeichnet, dass** die Selektion der Frequenz bzw. des Frequenzbandes mit der größten Signalenergie innerhalb der Frequenzgruppe vor Schritt (c) bzw. vor Schritt (d) erfolgt. 10
7. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** in Schritt (b) das Frequenzspektrum des abgezweigten TK-Signals nur in einem vorgegebenen Frequenzbereich erzeugt wird. 15
8. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** in Schritt (c) ein Frequenzspektrum zwischengespeichert wird, das durch eine Mittelung des aktuell in Schritt (b) erzeugten Frequenzspektrums mit vorher erzeugten Frequenzspektren gewonnen wird. 20
9. Verfahren nach Anspruch 8, **dadurch gekennzeichnet, dass** die Mittelung mit unterschiedlicher relativer Wichtung des aktuell erzeugten Frequenzspektrums in unterschiedlichen Frequenzbereichen erfolgt. 25
10. Verfahren nach Anspruch 9, **dadurch gekennzeichnet, dass** die Wichtung nach Kriterien der Psychoakustik gemäß den Mittelwerten des Wahrnehmungsspektrums des menschlichen Gehörs erfolgt. 30
11. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** in Schritt (e) ein nach vorgegebenen Kriterien mit einem Wichtungsfaktor $a < 1$ gewichtetes nachgebildetes Geräuschsignal vom aktuell ankommenden TK-Signal abgezogen wird. 35
12. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** dem in Schritt (d) erzeugten nachgebildeten Geräuschsignal vor Schritt (e) ein synthetisches Geräuschsignal beigemischt wird. 40
13. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das aktuell ankommende TK-Signal vor Schritt (e) einer definierten Zeitverzögerung 45

unterworfen wird, die vorzugsweise so ausgelegt ist, **dass** die Phasenlage des ankommenden TK-Signals mit der Phasenlage des nachgebildeten Geräuschsignals vor dem Abzug übereinstimmt.

14. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das aktuell ankommende TK-Signal unverzüglich dem Abzug in Schritt (e) zugeführt wird, und **dass** das nachgebildete Geräuschsignal in seiner Phasenlage vor Schritt (e) an die Phasenlage des aktuell ankommenden TK-Signals angepasst wird. 50

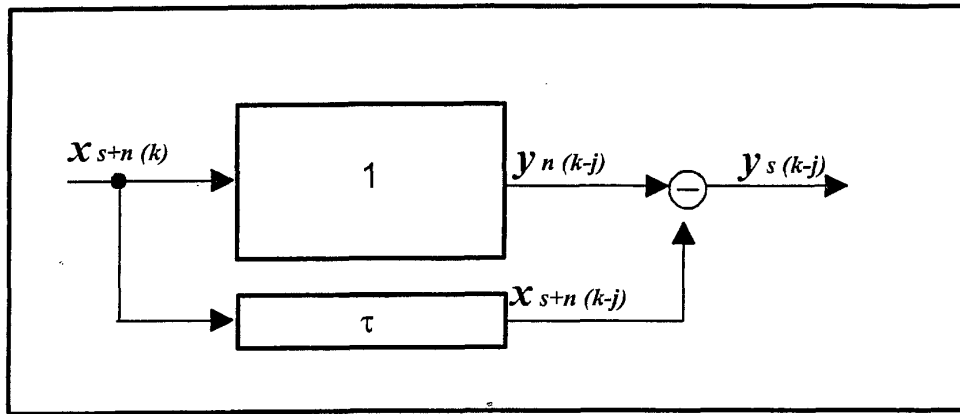


Fig1

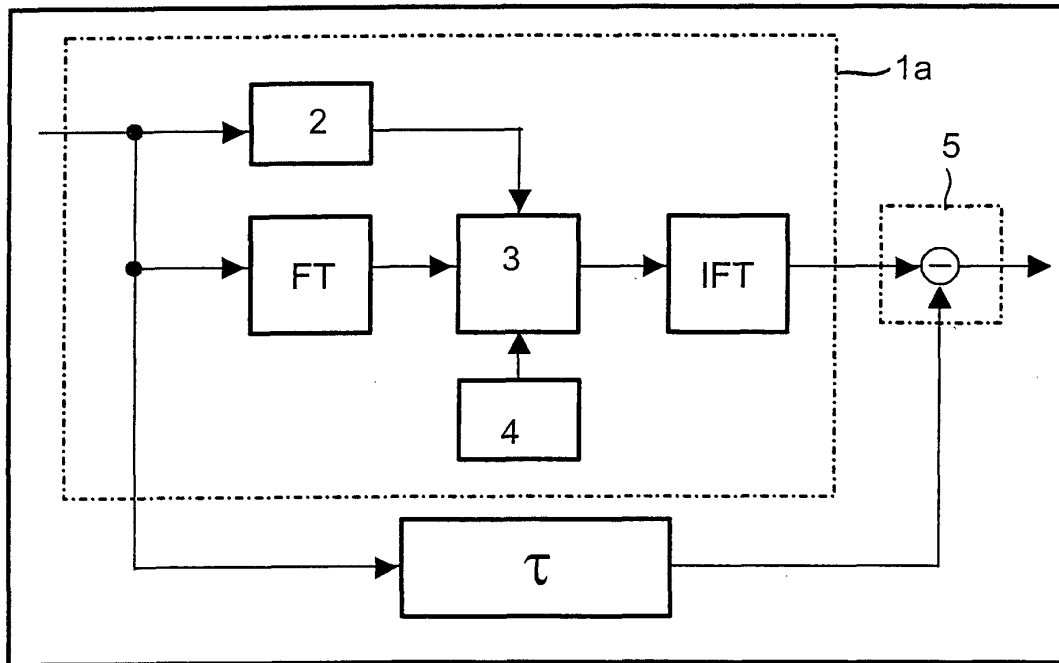


Fig.2

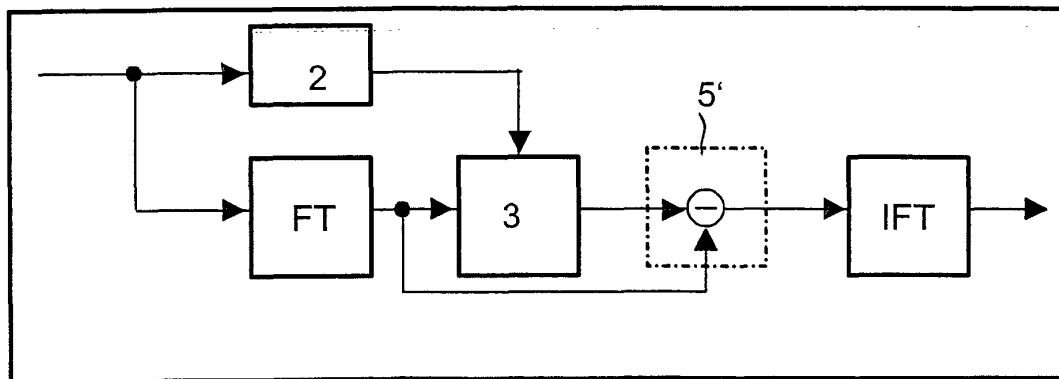


Fig.3

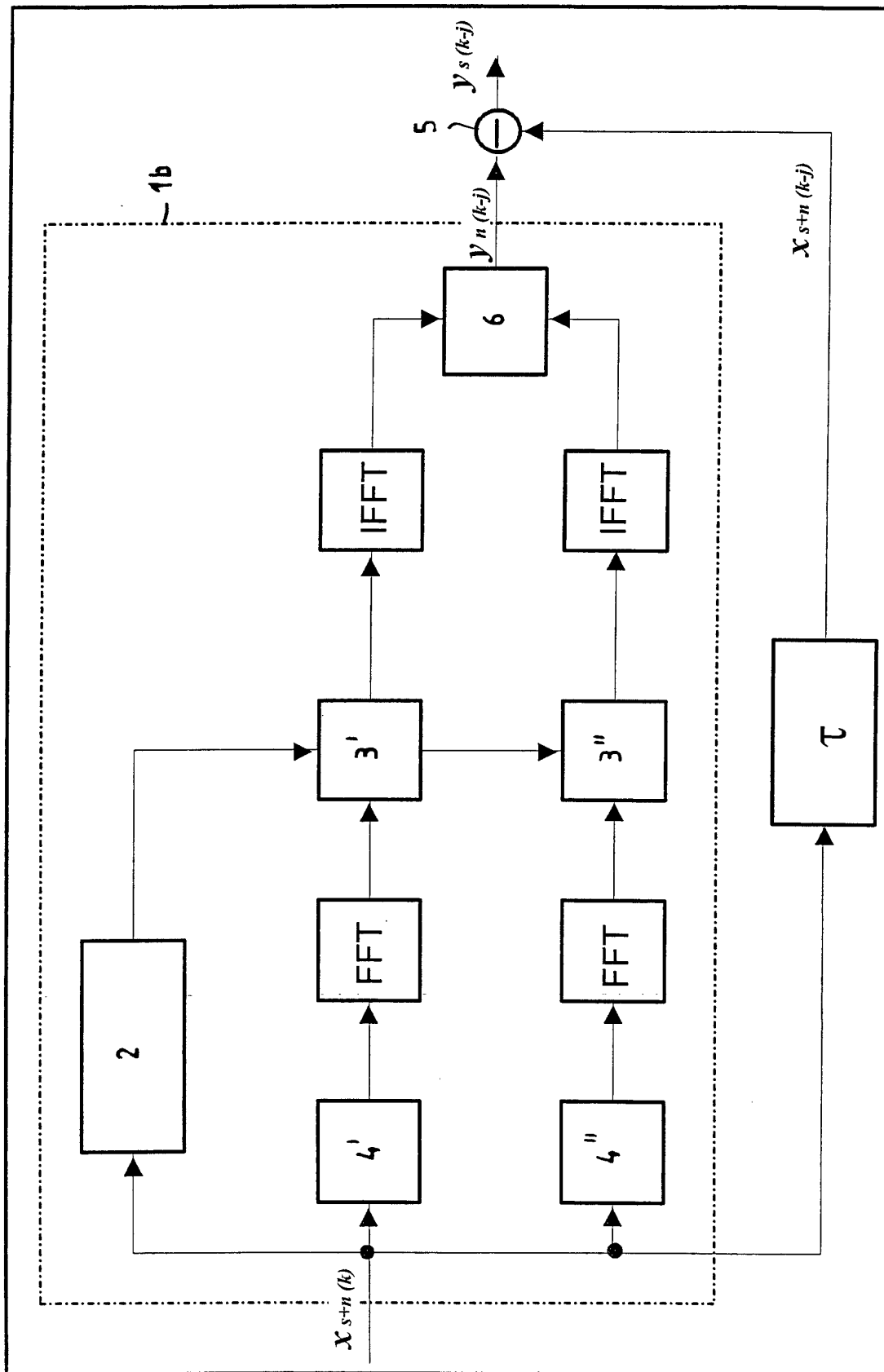


Fig.4

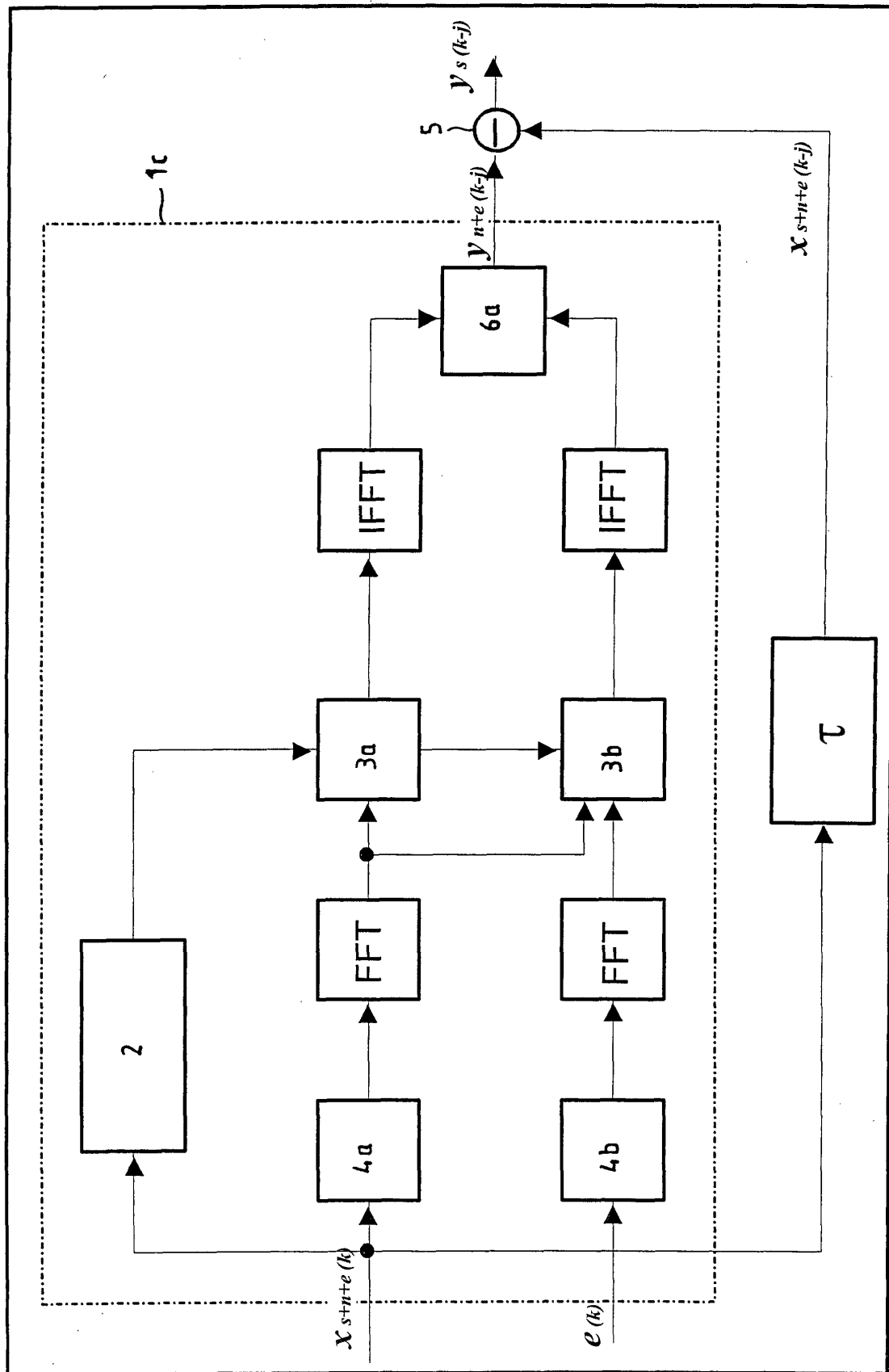


Fig.5

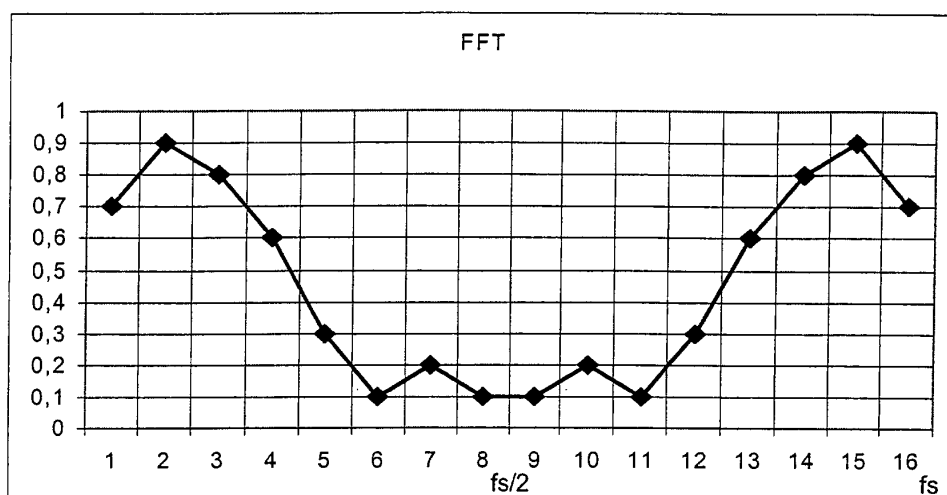


Fig.6a

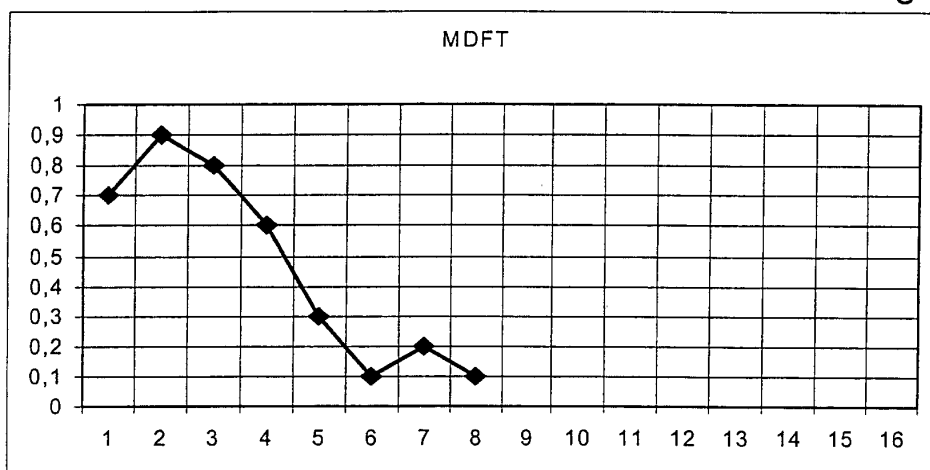


Fig.6b

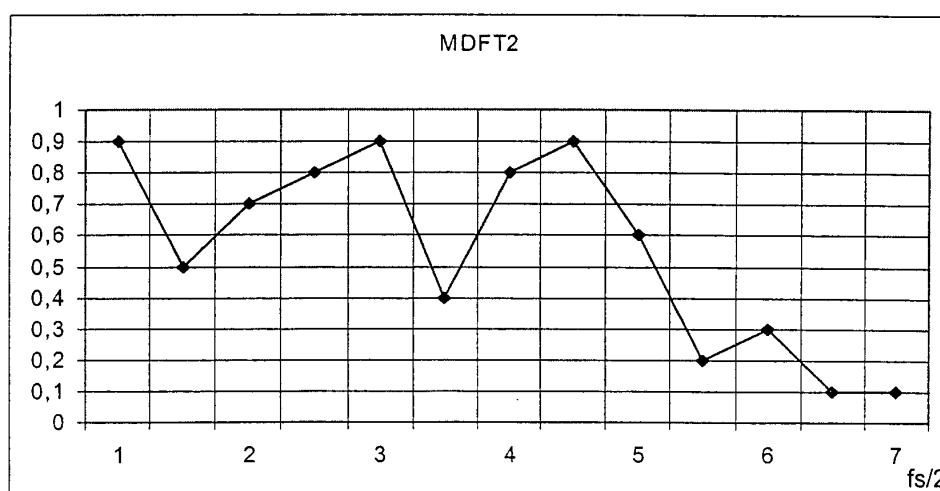


Fig.6c