



(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:
20.11.2002 Patentblatt 2002/47

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

(21) Anmeldenummer: **02005495.3**

(22) Anmeldetag: **11.03.2002**

(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: **Vierthaler, Matthias**
79108 Freiburg (DE)

(74) Vertreter: **Patentanwälte**
Westphal, Mussnug & Partner
Waldstrasse 33
78048 Villingen-Schwenningen (DE)

(30) Priorität: **18.05.2001 DE 10124699**

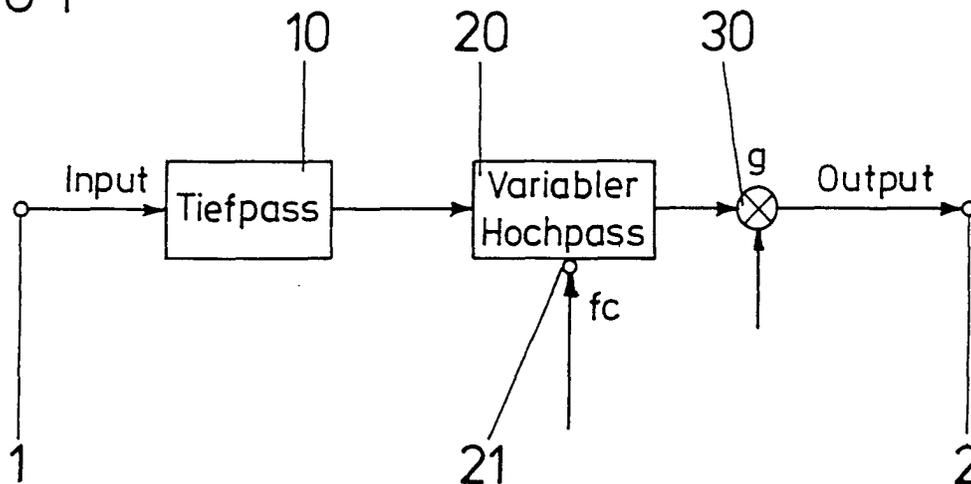
(71) Anmelder: **Micronas GmbH**
79108 Freiburg (DE)

(54) **Schaltungsanordnung zur Verbesserung der Verständlichkeit von Sprache enthaltenden Audiosignalen**

(57) Die Sprachverständlichkeit eines Audiosignal bei unveränderter Lautstärke wird verbessert, in dem das gesamte Audiosignal um einen konstanten Faktor angehoben und die Amplitude dieses angehobenen Si-

gnales durch einen Hochpass abgesenkt wird. Die Eckfrequenz f_c des Hochpasses wird dabei so geregelt, dass die Ausgangsamplitude des Audiosignals am Ende der Verarbeitungsstrecke gleich oder proportional der Eingangsamplitude des Audiosignals ist.

FIG 1



Beschreibung

[0001] Die Erfindung betrifft eine Schaltungsanordnung zur Verbesserung der Verständlichkeit von Sprache enthaltenden Audiosignalen gemäß den Merkmalen des Oberbegriffs des Anspruchs 1.

[0002] Es gibt verschiedene Möglichkeiten, wie die Sprachverständlichkeit von Audiosignalen verbessert werden kann. Eine Möglichkeit liegt in der Verbesserung des verrauschten Signals. Eine andere Möglichkeit liegt darin, solche Signale zu verbessern, die durch Hall und Echos etc. degradiert wurden. Schließlich kann ein gutes Audiosignal verändert werden, so dass es für Schwerhörige besser verständlich wird. Dies wird beispielsweise mit Hörgeräten erreicht. Letzlich ist die Veränderung eines guten Audiosignals möglich, so dass es bei starken Hintergrundgeräuschen besser verständlich ist.

[0003] Ziel der vorliegenden Erfindung ist es, die Sprachverständlichkeit eines verhältnismäßig guten Audiosignals bei unveränderter Lautstärke zu verbessern. Dies bedeutet, gleiche Verständlichkeit bei geringerer Lautstärke oder verbesserte Verständlichkeit bei Umgebungslärm.

[0004] Aus US 5,459,813 ist es bekannt, dass sogenannte "unvoiced sounds" (z. B. Konsonanten) von den viel stärkeren "voiced sounds" (z. B. Vokale) überdeckt werden. Da die "unvoiced sounds" wichtig für die Sprachverständlichkeit sind, wird in dieser Veröffentlichung vorgeschlagen, diese z. B. durch Clipping oder Amplitudenkompression zu verstärken.

[0005] In der Veröffentlichung "effects of amplitud distortion upon intelligibility of speech" von J. C. Liplider in dem Journal of acustical society of america, Oktober 1946 ist ein sogenanntes "peak clipping" bekannt. Ein solches "peak clipping" ohne Umgebungsrauschen hat kaum Einfluss auf die Sprachverständlichkeit. Ein "peak clipping" bei -20 dB führt immer noch zu einer Verständlichkeit von etwa 96%. Das sogenannte "center clipping" ist wesentlich schlechter, da hier die Konsonanten entfernt werden, die für die Verständlichkeit besonders wichtig sind. "Peak clipping" bei -24 dB braucht nur eine Verstärkung von etwa 14 dB, um dieselbe Verständlichkeit zu erreichen. Aus der Veröffentlichung Elwood Kretsinger et al "The Use of fast Limiting to improve the Intelligibility of Speech in Noise", Speech Monographs, March 1960 ist es bekannt, dass Konsonanten ca. 12 dB schwächer als Vokale sind. Verstärkt man die Konsonanten relativ zu den Vokalen, wird deshalb die Verständlichkeit von Sprache im Audiosignal erhöht. Ersetzt man den Clipper durch einen schnellen "peak limiter" (22 msec) kann man die Verständlichkeit noch weiter erhöhen. Bei -10 dB limiting erhöhte sich die Verständlichkeit von 56 % auf 84 %.

[0006] Aus Veröffentlichung Ian Thomas et al. "The Intelligibility of filtered-clipped Speech in Noise", The Journal of the Audio Engineering Society, June 1970 ist es bekannt, dass die Grundwelle eines Audiosignals,

das Sprache enthält, nur wenig zur Sprachverständlichkeit beiträgt, während die erste Resonanzfrequenz sehr wichtig ist. Deshalb sollte das Signal vor dem Clipping hochpassgefiltert werden.

[0007] Aus Veröffentlichung Ian Thomas et al., "Intelligibility enhancement through spectral weighting", Proceedings of the 1972 IEEE Conference on Speech Communication and Processing ist es bekannt, dass das Clipping zwar die Verständlichkeit von Sprache erhöht, jedoch die Signalqualität beeinträchtigt. In dieser Veröffentlichung wird deshalb vorgeschlagen, die Signalenergie in die signifikanten Frequenzbereiche zu verlagern.

[0008] Aus US 5,479,560 ist es darüber hinaus bekannt, das Audiosignal in mehrere Frequenzbänder aufzuteilen und diejenigen Frequenzbänder mit großer Energie verhältnismäßig stark zu verstärken und die anderen abzusenken. Dies wird deshalb vorgeschlagen, weil Sprache aus einer Aneinanderreihung von Phonemen besteht. Phoneme bestehen aus einer Vielzahl von Frequenzen. Diese werden an den Resonanzfrequenzen des Mund- und Rachenraums besonders verstärkt. Ein Frequenzband mit solche einem spektralen Peak wird Formant genannt. Formants sind besonders wichtig zur Erkennung von Phonemen und somit Sprache. Ein Ansatz zur Verbesserung der Sprachverständlichkeit ist es daher, die Peaks (Formants) des Frequenzspektrums eines Audiosignals zu verstärken und die dazwischen liegenden Täler abzuschwächen. Für einen Erwachsenen Mann liegt die Grundfrequenz von Sprache bei etwa 60 bis 250 Hz. Die ersten vier Formants liegen bei 500 Hz, 1 500 Hz, 2 500 Hz und 3 500 Hz (vgl. hierzu US-Patent 5,459,813).

[0009] Aus US 4,454,609 ist es bekannt, hauptsächlich die Konsonanten zu verstärken.

[0010] Schließlich beschreibt US 5,553,151 ein sogenanntes "forward masking". Hierbei werden schwache Konsonanten durch die vorhergehenden starken Vokale zeitlich überdeckt. Diese Veröffentlichung schlägt einen verhältnismäßig schnellen Kompressor mit einer "attack time" von ca. 10 msec. und einer "release time" von ca. 75 bis 150 msec. vor.

[0011] Problematisch bei den bisher bekannten Systemen zur Erhöhung der Sprachverständlichkeit von Sprache in Audiosignalen ist deren verhältnismäßig hohe Komplexität, das bedeutet, dass sowohl ein hoher Softwareaufwand zur Berechnung der einzelnen Algorithmen sowie ein hoher Hardwareaufwand notwendig ist. Bei einfacheren Systemen wird dagegen das Audiosignal so verändert, dass die Sprache nicht mehr sehr natürlich klingt. Des Weiteren kann bei einfachen Systemen dem Sprachsignal Störungen zugefügt werden, das einer verbesserten Verständlichkeit sogar entgegen wirken kann.

[0012] Ziel der vorliegenden Erfindung ist es daher, eine Schaltungsanordnung zur Verbesserung der Sprachqualität von Audiosignalen anzugeben, das einerseits geringen Aufwand erfordert und andererseits

die Sprache noch natürlich klingen lässt.

[0013] Dieses Ziel wird durch eine Schaltungsanordnung mit dem Merkmale des Anspruchs 1 gelöst.

[0014] Weiterbildungen einer solchen Schaltungsanordnung sind Gegenstand der Unteransprüche.

[0015] Die Erfindung beruht im Wesentlichen darauf, das Audiosignal auf einen vorgegebenen Faktor zu verstärken und in einem Hochpass zu filtern, wobei die Eckfrequenz des Hochpasses so geregelt wird, dass die Amplitude des Audiosignals nach der Verarbeitungsstrecke gleich oder proportional der Amplitude des Audiosignals am Eingang der Verarbeitungsstrecke ist.

[0016] Mit dieser Schaltungsanordnung kann die Grundwelle des Sprachsignals, die relativ wenig zur Verständlichkeit beiträgt, aber die größte Energie besitzt, abgeschwächt werden und das übliche Signalspektrums des Audiosignals entsprechend angehoben werden. Außerdem kann die Amplitude der Vokale (große Amplitude, tiefe Frequenz) im Übergangsbereich Konsonant (kleine Amplitude, große Frequenz) zu Vokal abgesenkt werden, um das sogenannte "backward masking" zu verringern. Dazu wird das gesamte Signal um einen Faktor g angehoben. Dieser Faktor steuert die Stärke des Effekts der Signalverbesserung, wobei sinnvolle Werte für den Faktor g etwa zwischen 1,5 und 4 liegen. Mit der erfindungsgemäßen Schaltungsanordnung werden als höher frequente Anteile angehoben und die tieffrequente Grundwelle im gleichen Maße abgesenkt, so dass die Amplitude (oder Energie) des Audiosignales unverändert bleibt. Für Signalanteile mit kleinen Amplituden, also Konsonanten, kann mit der Schaltungsanordnung nach der vorliegenden Erfindung die Eckfrequenz des variablen Hochpasses abgesenkt werden. Deshalb kann in der Regelung zu dem Eingangssignal noch ein "offset" addiert werden, der entweder fix oder proportional zur Peak-Amplitude des eingangsseitigen Audiosignals ist.

[0017] In einer Weiterbildung der Erfindung ist vorgesehen, dass höherfrequente Signalanteile im Audiosignal abgesenkt werden. Mit einem Tiefpass vor dem variablen Hochpass können Störungen im Signal unterdrückt werden.

[0018] In einer Weiterbildung der Erfindung ist vorgesehen, dass die Eckfrequenz f_c des variablen Hochpassfilters nach unten begrenzt wird, da die unterste Frequenz für Sprache bei ca. 200 Hz liegt. Bewährt hat sich für eine untere Eckfrequenz ein Bereich von etwa 100 bis 120 Hz.

[0019] Nachfolgend wird die erfindungsgemäße Schaltungsanordnung anhand von Figuren beispielhaft erläutert. Es zeigen:

Figur 1 die prinzipielle Schaltungsanordnung zur Verbesserung der Sprachverständlichkeit in einem Audiosignal,

Figur 2 eine Weiterbildung der Schaltungsanordnung von Figur 1,

Figur 3 eine andere Weiterbildung der Schaltungsanordnung von Figur 1, und

5 Figur 4 eine andere Weiterbildung der Schaltungsanordnung von Figur 1, und

Figur 5 eine vierte Weiterbildung der erfindungsgemäßen Schaltungsanordnung.

10 **[0020]** In den nachfolgenden Figuren bezeichnen gleiche Bezugszeichen, sofern nicht anders angegeben, gleiche Teile mit gleicher Bedeutung.

[0021] In Figur 1 ist der prinzipielle Aufbau der erfindungsgemäßen Schaltungsanordnung gezeigt. Die Schaltungsanordnung weist einen variablen Hochpass 20 auf, der in seiner Eckfrequenz f_c veränderbar ist. Hierfür verfügt der variable Hochpass 20 über einen Steuereingang 21, an dem ein Steuersignal zur Veränderung der Eckfrequenz f_c anlegbar ist. Diesem variablen Hochpass 20 wird vorzugsweise über einen Tiefpass 10 das zu verbessernde Audiosignal zugeführt. Hierfür ist eine Eingangsklemme 1 zum Anlegen des Audiosignals vorgesehen. Der Tiefpass 10 muss nicht vorgesehen sein, ist jedoch vorteilhaft, um Signalstörungen im Audiosignal zu beseitigen. Am Ausgang des variablen Hochpasses 20 sitzt eine Verstärkerstufe 30, die das ausgangsseitige Signal des variablen Hochpasses 20 um einen Faktor g verstärkt. Dieser Faktor g ist einstellbar und liegt vorzugsweise zwischen etwa 1,5 und 4. Ein einmal eingestellter Verstärkungsfaktor wird vorzugsweise nicht mehr verändert. Die gesamte Verarbeitungsstrecke bestehend aus variablen Hochpass 20 und Verstärker 30 sowie optionalem Tiefpass 10 verfügt über eine Ausgangsklemme 2, an der das verarbeitete Audiosignal als Ausgangssignal abgreifbar ist.

[0022] Erfindungsgemäß wird eine Regelung der Eckfrequenz f_c des variablen Hochpasses 20 in folgender Art und Weise zur Verbesserung der Sprachverständlichkeit von Sprache innerhalb des Audiosignals durchgeführt. Ist die Amplitude (oder auch Energie) des Eingangssignals am Eingang 1 der Schaltungsanordnung größer als die Amplitude (oder Energie) am Ausgang 2 der Übertragungsstrecke, dann wird die Eckfrequenz f_c erniedrigt. Im Übrigen erhöht. Sofern die Amplituden am Eingang 1 und Ausgang 2 gleich oder zu einem vorgegebenen Faktor proportional sind, erfolgt keine weitere Veränderung der Eckfrequenz f_c .

[0023] In Figur 2 ist eine Weiterbildung der Schaltungsanordnung von Figur 1 dargestellt. In Figur 2 ist ein Vergleichler 36 mit nachgeschaltetem Integrator, dem ein Skalierungsfaktor K_i vorgeschaltet ist, vorgesehen. Die Ausgangsklemme des Integrators 40 ist mit dem Steuereingang 21 des variablen Hochpasses 20 in Verbindung. Der Vergleichler 36 weist zwei Eingangsklemmen 34, 35 auf, an deren erste Klemme 34 das Eingangssignal und an dessen Klemme 35 das Ausgangssignal der Übertragungsstrecke angelegt wird.

[0024] Die Schaltungsanordnung von Figur 3 unter-

scheidet sich von der Schaltungsanordnung von Figur 2 dadurch, dass der Integrator 40 durch eine digitale Schaltungsanordnung 60 ersetzt ist. In der digitalen Schaltungsanordnung 60 wird nach Maßgabe des Ausgangssignals des Vergleichers 36 die Eckfrequenz f_c um einen Schritt d erhöht oder erniedrigt, je nachdem, ob das Ausgangssignal x_c am Ausgang des Vergleichers 36 größer oder kleiner 0 ist.

[0025] Schließlich ist in Figur 4 noch eine Weiterbildung der erfindungsgemäßen Schaltungsanordnung dargestellt. Die Weiterbildung besteht darin, dass zu dem an dem Eingang 34 anstehenden Eingangssignal ein Offset K addiert wird. Dieser Offset kann konstant gewählt werden oder ein mit einem Faktor K gewichteter Ausgang eines Peak-Detektor 70 sein. An dem Peak-Detektor 70 wird eingangsseitig das Audiosignal angelegt.

[0026] Mit der erfindungsgemäßen Schaltungsanordnung gemäß den Figuren 1 bis 4 ist es möglich, die Grundwelle des Audiosignals abzusenken und den restlichen Signalanteil anzuheben. Hierfür ist das variable Hochpassfilter 20 verantwortlich.

[0027] Für den Fall, dass im Sprachsignal ein Konsonant einem Vokal folgt, arbeitet die Schaltungsanordnung folgendermaßen: Ein Vokal ist tieffrequent mit großer Amplitude. Ein Konsonant ist dagegen hochfrequent mit kleiner Amplitude. Bei der erfindungsgemäßen Schaltungsanordnung wird der Verstärkungsfaktor g so eingestellt, dass eine Verstärkung von 6 dB erreicht wird. Durch den tieffrequenten Vokal hat sich die Eckfrequenz des variablen Hochpassfilters 20 auf diese tiefe Frequenz eingestellt. Die Grundwelle ist also so weit abgesenkt, dass die Ausgangsamplitude gleicher Eingangsamplitude des Audiosignals ist, obwohl die Verstärkung von 6 dB gewählt wurde. Folgt auf den Vokal nun ein Konsonant (höhere Frequenz!) wird dieser sofort um 6 dB angehoben, da die Eckfrequenz des Hochpassfilters 20 noch auf die tiefe Frequenz des Vokals eingestellt ist. Der Konsonant wird als weniger stark vom Vokal überdeckt. Erst nach einigen Millisekunden erhöht sich die Eckfrequenz f_c und senkt somit auch den Konsonant ab, so dass die Amplitude des Eingangssignals gleich der Amplitude des Ausgangssignals der Verarbeitungstrecke ist.

[0028] Bei einem Übergang Konsonant auf Vokal arbeitet die erfindungsgemäße Schaltungsanordnung von Figur 1 folgendermaßen. Das Hochpassfilter 20 hat sich auf die Frequenz des Konsonants eingestellt. Die Amplitude des Eingangssignals entspricht der Amplitude des Ausgangssignals. Folgt nun ein Vokal (tieffrequent) wird durch die verhältnismäßig hohe Eckfrequenz f_c des Hochpassfilters 20 der Vokal beim zeitlichen Übergang gedämpft und der Konsonant folglich nicht überdeckt. Erst nach einigen Millisekunden ist die Eckfrequenz f_c aufgrund der Regelzeit der Regelschleife so eingeregelt, dass die Amplitude des Eingangssignals der Amplitude des Ausgangssignals entspricht.

[0029] Abschließend ist noch folgendes anzumerken:

Bei einem Stereosignal kann entweder jeder Kanal eine eigene Regelung erhalten wie oben beschrieben oder sie können eine gemeinsame Regelung benutzen. Dann ist z. B. (vgl. Figur 5) an den Eingang $34 = \text{Abs}(\text{Input_Left}) + \text{Abs}(\text{Input_Right})$ anzulegen und an den Eingang $35 = \text{Abs}(\text{Output_Left}) + \text{Abs}(\text{Output_Right})$. Der Audiopfad (Hochpass, Tiefpass, Gain) wird für links und rechts getrennt berechnet, die Hochpässe besitzen aber dieselbe Eckfrequenz f_c .

Patentansprüche

1. Schaltungsanordnung zur Verbesserung der Verständlichkeit von Sprache enthaltenden Audiosignalen, bei welcher Frequenz- und/oder Amplitudenanteile des Audiosignals nach vorgegebenen Parametern verändert werden, **dadurch gekennzeichnet, dass** das Audiosignal in einer Verarbeitungstrecke um einen vorgegebenen Faktor g verstärkt sowie in einem Hochpasses (20) geführt wird, wobei eine Eckfrequenz f_c des Hochpasses (20) so regelbar ist, dass die Amplitude des Audiosignals (2) nach der Verarbeitungstrecke gleich oder proportional der Amplitude des Audiosignals vor der Verarbeitungstrecke ist.
2. Schaltungsanordnung nach Anspruch 1, **dadurch gekennzeichnet, dass** der Faktor $g > = 1$ gewählt ist.
3. Schaltungsanordnung nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** der Faktor g etwa zwischen 1,5 und 4 gewählt ist.
4. Schaltungsanordnung nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet, dass** dann, wenn die Amplitude des Eingangssignals größer als die Amplitude des Ausgangssignals am Ausgang der Verarbeitungstrecke ist, die Eckfrequenz f_c erniedrigt wird und im umgekehrten Fall erhöht wird.
5. Schaltungsanordnung nach einem der Ansprüche 1 bis 4, **dadurch gekennzeichnet, dass** die Änderung der Eckfrequenz f_c inkremental erfolgt, vorzugsweise in ein Hz-Schritten.
6. Schaltungsanordnung nach einem der Ansprüche 1 bis 5, **dadurch gekennzeichnet, dass** die Eckfrequenz f_c im Bereich von etwa 100 Hz bis 1 kHz variierbar ist.
7. Schaltungsanordnung nach einem der Ansprüche 1 bis 6, **dadurch gekennzeichnet, dass** die untere Eckfre-

quenz f_c bei etwa 100 bis 120 Hz liegt.

8. Schaltungsanordnung nach einem der Ansprüche 1 bis 7,
dadurch gekennzeichnet, dass vor dem variablen Hochpass (20) ein Tiefpass (10) geschaltet ist. 5
9. Schaltungsanordnung nach Anspruch 8,
dadurch gekennzeichnet, dass der Tiefpass (10) eine Eckfrequenz bei etwa 6 kHz aufweist. 10
10. Schaltungsanordnung nach einem der Ansprüche 1 bis 9,
dadurch gekennzeichnet, dass an einen Steuereingang (21) des variablen Hochpasses (20) zur Veränderung der Eckfrequenz (f_c) ein Vergleichs- (36) gekoppelt ist, an dessen einen Eingang (34) das Eingangssignal der Verarbeitungstrecke und an dessen anderen Eingang (35) das Ausgangssignal der Verarbeitungstrecke geschaltet ist. 15
 20
11. Schaltungsanordnung nach Anspruch 10,
dadurch gekennzeichnet, dass zwischen den Steuereingang (21) des variablen Hochpasses (20) und den Ausgang des Vergleichers (36) ein Integrator (40) geschaltet ist. 25
12. Schaltungsanordnung nach Anspruch 10,
dadurch gekennzeichnet, dass zwischen den Steuereingagn (21) des variablen Hochpasses (20) und den Ausgang des Vergleichers (36) eine digitale Schaltungsanordnung (60) zum Inkrementieren der Eckfrequenz f_c in Schritten (d) vorgesehen ist. 30
 35
13. Schaltungsanordnung nach einem der Ansprüche 10 bis 12,
dadurch gekennzeichnet, dass dem Eingangssignal an einem Eingang (34) des Vergleichers (36) ein Offset addiert wird. 40
14. Schaltungsanordnung nach einem der Ansprüche 10 bis 13,
dadurch gekennzeichnet, dass das Audiosignal ein Stereosignal ist, und dass einem ersten Eingang (34) des Vergleichers (36) die Summe aus den Eingangssignalen für den linken und rechten Kanal und dass dem zweiten Eingang (35) des Vergleichers (36) die Summe aus dem Ausgangssignal für den linken und rechten Kanal zugeführt wird. 45
 50

55

FIG 1

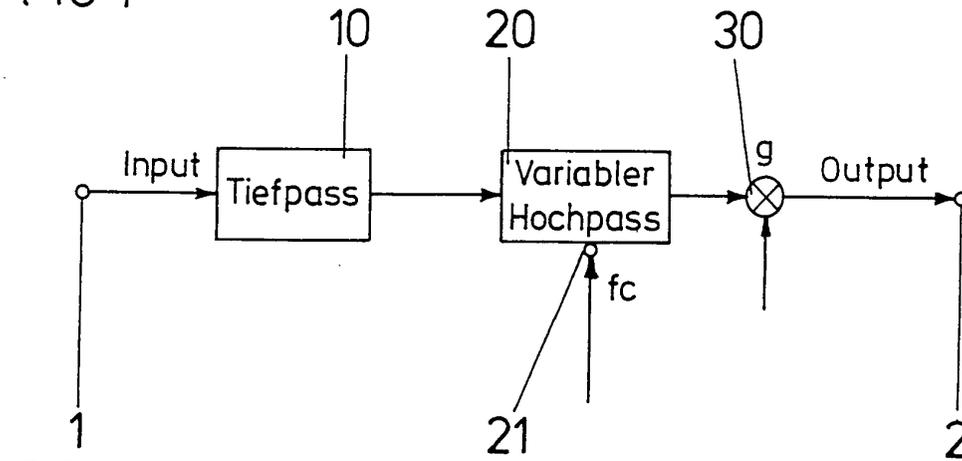


FIG 2

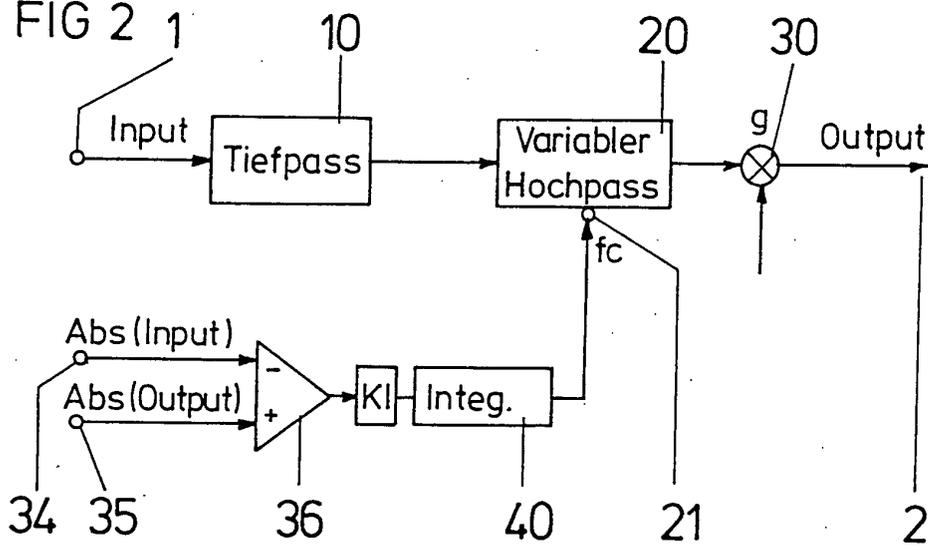


FIG 3

