

METODO PER MIGLIORARE LA PERCEZIONE DELLA QUALITA' DI UN SEGNALE AUDIO  
DIGITALE EMESSO DA UN RICEVITORE DI SEGNALI TELEVISIVI, IN PARTICOLARE DEL

*Original*

METODO PER MIGLIORARE LA PERCEZIONE DELLA QUALITA' DI UN SEGNALE AUDIO  
DIGITALE EMESSO DA UN RICEVITORE DI SEGNALI TELEVISIVI, IN PARTICOLARE DEL  
TIPO A SCHERMO PIATTO, E RELATIVO DISPOSITIVO / Scopece, Leonardo; Carullo, Alessio; Masoero, Marco Carlo;  
Shtrepi, Louena; Riente, Fabrizio; Astolfi, Arianna; Rocchi, Luigi; Fatale, Giorgio. - (2020).

*Availability:*

This version is available at: 11583/2966216 since: 2022-07-08T14:51:33Z

*Publisher:*

*Published*

DOI:

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

Descrizione dell'Invenzione Industriale dal titolo: **-RRI031-**  
**“METODO PER MIGLIORARE LA PERCEZIONE DELLA QUALITÀ DI UN  
SEGNALE AUDIO DIGITALE EMESSO DA UN RICEVITORE DI SEGNALI  
TELEVISIVI, IN PARTICOLARE DEL TIPO A SCHERMO PIATTO, E  
RELATIVO DISPOSITIVO”**

a nome RAI Radiotelevisione Italiana S.p.A., di nazionalità italiana, con sede in Viale Mazzini 14, 00195 ROMA -P.IVA: 06382641006 ed elettivamente domiciliata, ai fini del presente incarico, presso i Mandatari CAMOLESE Marco (No. Iscr. Albo 882BM), BARONI Matteo (No. Iscr. Albo 1064BM), BIANCO Mirco (No. Iscr. Albo 1639B), BORSANO Corrado (No. Iscr. Albo 446BM), FERRONI Filippo (No. Iscr. Albo 530BM) e REPOSIO Giancarlo (No. Iscr. Albo 1168BM) c/o Metroconsult S.r.l., Via Sestriere 100, 10060 NONE (TO).

Inventori designati:

- SCOPECE Leonardo, Via Gallina 4, 10092 Beinasco (TO)
- FATALE Giorgio, Via Vallemaio 23, 00189 Roma (RM)
- ROCCHI Luigi, Via Misurina 71, 00135 Roma (RM)
- ASTOLFI Arianna, Corso San Maurizio 53, 10124 Torino (TO)
- RIENTE Fabrizio, Via Barme Solan 42, 11029 Verres (AO)
- SHTREPI Louena, Via Staffarda 15, 10141 Torino (TO)
- MASOERO Marco, Via Giolitti 49 bis, 10123 Torino (TO)
- CARULLO Alessio, Via G. Perotti 10, 10015 Ivrea (TO)

Depositata il

No.

\*\*\*\*\*

### **DESCRIZIONE**

La presente invenzione si riferisce ad un metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi e da un relativo dispositivo.

Più in particolare, la presente invenzione si riferisce ad un metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore televisivo del tipo a schermo piatto, e relativo dispositivo.

In particolare, il miglioramento della qualità del segnale audio percepita è relativo al parlato, di cui si migliora la chiarezza e di conseguenza l'intelligibilità. La presente invenzione ha come base scientifica la rilevanza che alcune frequenze dello spettro della voce hanno ai fini dell'intelligibilità del parlato: si vedano a questo proposito lo Standard ANSI S3.5-1997 e l'articolo a nome Steeneken e Hougast: "Mutual Dependence of the octave-band weights in predicting speech intelligibility", *Speech Communication* 28, pagine 109-123.

Le frequenze di maggiore interesse ai fini di una buona intelligibilità sono quelle comprese fra i 500 Hz e i 4 kHz.

La Figura 1 illustra una tabella in cui sono evidenziati i fattori  $a_f$  che rappresentano il contributo relativo di ogni banda di ottava all'indice di intelligibilità Speech Transmission Index (STI) in accordo con lo Standard IEC 60268-11.

L'indice STI è una misura di qualità di trasmissione del parlato.

La Figura 2 illustra i fattori di pesatura della banda per il contributo di intelligibilità, in bande di ottava, relativi a diverse tipologie di test di intelligibilità, riportati nel succitato articolo.

La Figura 3 illustra i fattori di pesatura della banda per il contributo di intelligibilità, in bande di ottava, relativi a diversi test di intelligibilità, riportati nello Standard ANSI S3.5-1997.

La Figura 4 illustra i fattori di pesatura della banda per il contributo di intelligibilità, in bande di un terzo di ottava, relativi a diversi test di intelligibilità, riportati nello Standard ANSI S3.5-1997.

In generale, il parlato subisce continue fluttuazioni di livello connesse al linguaggio. La dinamica della voce nel caso di conversazione a voce normale è di circa 30 dB, intesa come differenza tra il valore massimo e il valore minimo del livello emesso, pause di silenzio escluse. I suoni vocalici sono più forti e a frequenza più bassa, e durano più a lungo dei suoni delle consonanti, che possono essere anche mute o avere carattere impulsivo. I livelli massimi sono attribuibili dunque ai suoni vocalici, mentre i livelli minimi a certe consonanti sonore, che maggiormente contribuiscono alla intelligibilità del messaggio.

Lo sviluppo di ricevitori di segnali televisivi con schermi sempre più sottili, accompagnato da dettami stilistici volti alla ricerca di linee minimaliste, ha portato i costruttori di ricevitori di segnali televisivi ad adottare altoparlanti sempre più miniaturizzati e che irradiano verso il basso o il retro del ricevitore, a scapito della qualità audio percepita dall'utente.

Ciò si verifica in particolare quando i ricevitori di segnali televisivi sono del tipo a schermo piatto, anche noti come "Flat TV".

Scopo della presente invenzione è pertanto quello di indicare un metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi, in particolare del tipo a schermo piatto, e relativo dispositivo, che migliori la qualità audio, in particolare del parlato, percepita da un utente di tale ricevitore.

Un ulteriore scopo della presente invenzione è quello di indicare un metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi, in particolare del tipo a schermo piatto, e relativo dispositivo, che possa apportare miglioramenti alla qualità del segnale audio percepita da un utente indifferentemente dalla tipologia di contenuto veicolata dal segnale audio.

Un ulteriore scopo della presente invenzione è quello di indicare un metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi, in particolare del tipo a schermo piatto, e relativo dispositivo che possa apportare miglioramenti alla qualità del segnale audio percepita dall'utente indifferentemente dal fatto che il segnale televisivo sia trasmesso via televisione digitale terrestre, satellite o Internet.

In una fase preliminare della propria ricerca, la Richiedente ha sperimentalmente analizzato una pluralità di tracce audio appartenenti a tre tipologie di contenuti, ossia "Parlato", "Canto e Musica" e "Sport", effettivamente messi in onda. In particolare, le tracce audio del genere "Parlato" sono state estratte da fiction, film, documentari e news; quelle del genere "Canto e Musica" sono state estratte da programmi musicali; e infine quelle del genere "Sport" sono state estratte da eventi sportivi come ad esempio partite di calcio, pallavolo e basket.

L'analisi ha evidenziato come le tracce dei generi "Parlato" e "Sport" presentino sostanzialmente lo stesso spettro a tutte le frequenze, mentre le tracce del genere "Canto e Musica" si discostino dagli altri due generi soprattutto alle basse frequenze. Tuttavia, focalizzando l'interesse sulle frequenze più importanti per l'intelligibilità del parlato, ossia tra 500 Hz e 4 kHz, tutti e tre i generi risultano possedere uno spettro simile.

In una forma di realizzazione dell'invenzione, un dispositivo comprende un processore di segnale audio digitale opportunamente programmato per migliorare l'intelligibilità del parlato. Il ruolo del processore di segnale digitale consiste nell'alterare il segnale audio in ingresso al dispositivo e trasformarlo opportunamente prima di essere trasmesso via etere e raggiungere il ricevitore di segnali televisivi dell'utente.

Il trattamento audio applicato all'interno del dispositivo prevede una equalizzazione dinamica, cioè il segnale audio viene trattato in dipendenza dal contenuto in frequenza, e viene adattato in tempo reale ad una curva di riferimento in frequenza. In particolare, se un segnale audio in ingresso al dispositivo presenta già un livello di pressione sonora nell'intervallo di frequenze scelto per il miglioramento simile a quello della curva di riferimento, l'equalizzatore non forza modifiche in frequenza, perché altrimenti quest'ultimo provocherebbe una distorsione non trascurabile del segnale stesso.

Viceversa, se il livello di pressione sonora nell'intervallo è non conforme alla curva di riferimento, il processore di segnale digitale cercherà di far coincidere automaticamente lo spettro del segnale in ingresso alla curva di riferimento.

Il segnale audio trattato dal dispositivo non deve eccedere un certo livello di loudness, definita come la percezione soggettiva dell'intensità sonora, e una sua massima deviazione.

L'equalizzazione dinamica è stata programmata come segue:

- le frequenze sotto i 100 Hz non possono essere efficacemente riprodotte dagli altoparlanti interni del ricevitore di segnali televisivi e non sono state quindi trattate;
- le frequenze comprese nell'intervallo 0,1-1 kHz sono meno importanti per

- l'intelligibilità del parlato e non sono state quindi trattate;
- le frequenze comprese nell'intervallo 1-4 kHz sono le più importanti per l'intelligibilità del parlato e sono quindi state trattate;
  - le frequenze oltre i 6,3 kHz sono meno importanti per l'intelligibilità del parlato e non possono essere riprodotte in modo ottimale dagli altoparlanti del ricevitore di segnali televisivi ma sono state comunque trattate.

La presente invenzione prevede pertanto lo sviluppo di una curva di riferimento in frequenza, o funzione di trasferimento, implementata su un processore di segnale audio digitale all'interno di un dispositivo.

Il segnale in ingresso al dispositivo viene trattato dinamicamente in base alla suddetta curva di riferimento per compensare i fattori di degradazione che compromettono l'intelligibilità del parlato del contenuto audio.

Il dispositivo hardware può essere parte del ricevitore di segnali televisivi ma più generalmente viene installato nella catena di trasmissione del segnale televisivo in modo da essere ricevibile tramite le varie piattaforme diffusive (digitale terrestre, satellite, Internet).

Ulteriori caratteristiche vantaggiose della presente invenzione sono oggetto delle allegate rivendicazioni.

Tali caratteristiche ed ulteriori vantaggi della presente invenzione risulteranno maggiormente chiari dalla descrizione di un suo esempio di realizzazione mostrato nei disegni annessi, forniti a puro titolo esemplificativo e non limitativo, in cui:

- la Figura 1 illustra una tabella in cui sono evidenziati i fattori  $a_f$  che rappresentano il contributo relativo di ogni banda di ottava all'indice di intelligibilità Speech Transmission Index (STI) in accordo con lo Standard IEC 60268-11;
- la Figura 2 illustra i fattori di pesatura della banda per il contributo di intelligibilità, in bande di ottava, relativi a diverse tipologie di test di intelligibilità, riportati nel riferimento Steeneken e Hougast: "Mutual Dependence of the octave-band weights in predicting speech intelligibility", Speech Communication 28, pagine 109-123;

- la Figura 3 illustra i fattori di pesatura della banda per il contributo di intelligibilità, in bande di ottava, relativi a diversi test di intelligibilità, riportati nello Standard ANSI S3.5-1997;
- la Figura 4 illustra i fattori di pesatura della banda per il contributo di intelligibilità, in bande di un terzo di ottava, relativi a diversi test di intelligibilità, riportati nello Standard ANSI S3.5-1997;
- la Figura 5 illustra uno schema a blocchi di un dispositivo di ottimizzazione di un segnale audio secondo l'invenzione;
- la Figura 6 illustra alcuni moduli fondamentali di un processore di segnale audio digitale utilizzabile nel dispositivo di Figura 5;
- la Figura 7 illustra un diagramma di flusso di un metodo per migliorare la percezione della qualità di un segnale audio secondo la presente invenzione;
- le Figure 8 e 9 illustrano un grafico di una curva di riferimento di equalizzazione in frequenza implementata nel processore di segnale audio digitale di Figura 6;
- la Figura 10 rappresenta una tabella che indica per determinati valori di frequenza del grafico di Figure 8 e 9 la massima attenuazione/amplificazione applicabile dal processore di segnale audio digitale di Figure 5 e 6 in una determinata banda di frequenza;
- la Figura 11 rappresenta un ingrandimento della Figura 9 in corrispondenza dell'intervallo di frequenze 4-16 kHz;
- le Figure 12, 13 e 14 illustrano tre esempi di applicazione della curva di riferimento rispettivamente a contenuti audio di tre generi diversi, rispettivamente "Parlato", "Musica e Canto" e "Sport".

Con riferimento alla Figura 5, viene illustrato uno schema a blocchi di un dispositivo 1 di ottimizzazione, definito anche Digital Audio Optimizer (DAO), di un segnale audio/video digitale 3, in particolare un segnale SDI, secondo l'invenzione. Qualora il segnale audio/video digitale in ingresso al dispositivo 1 fosse di formato diverso, esso viene preventivamente convertito in formato SDI. Il dispositivo 1 può comprendere varie interfacce input/output, come ad esempio connessioni input/output di tipo BNC (Bayonet Neill Concelman) atte

al trasporto di segnali audio/video embedded SDI e HD-SDI, preferibilmente un ingresso e quattro uscite in BNC per SDI e HD-SDI, e una connessione di rete, in particolare di tipo LAN, per collegamenti verso "porte" Internet.

Il dispositivo 1 comprende mezzi di ingresso 4 del segnale audio/video digitale 3, in particolare una porta di ingresso di tipo SDI, e mezzi 9, in particolare un de-embedder, per estrarre un primo segnale audio 10 dal segnale audio/video digitale 3. Il de-embedder 9 è in grado di gestire vari formati audio quali, a titolo di esempio, PCM, Dolby E, Dolby Digital, Dolby Digital Plus, MPEG-4 HE AAC, MPEG-4 AAC e così via.

Il dispositivo 1 comprende un processore 5 di segnale audio digitale, ad esempio un processore Junger D\*AP4 FLX. All'interno del processore 5 di segnale audio digitale sono compresi mezzi di memoria 7 di tipo non volatile in cui è implementata una curva di riferimento 30, o funzione di trasferimento, la cui rilevanza verrà descritta nel seguito.

Il dispositivo 1 comprende inoltre mezzi 12, in particolare un de-embedder, per ricombinare un secondo segnale audio 11 in uscita dal processore 5 di segnale audio digitale con un segnale video 6 appartenente al segnale audio/video digitale 3, estratto e opportunamente ritardato da un blocco estrattore/ritardatore 13. Il secondo segnale audio 11 viene così ricombinato con il segnale video 6 ed esce dal dispositivo 1 attraverso una porta di uscita 15 di tipo SDI come terzo segnale audio digitale 17 trattato secondo il metodo oggetto della presente invenzione.

Il dispositivo 1 può prevedere un modulo di by-pass automatico (non illustrato) dell'applicazione del metodo oggetto della presente invenzione che si attiva in caso di avaria del processore 5 di segnale audio digitale oppure dell'intero dispositivo 1, in modo da ottenere in uscita il segnale audio/video digitale 3 applicato in ingresso al dispositivo 1 senza alterazioni e senza ritardi. Tale avaria può anche essere segnalata da mezzi di segnalazione, in particolare un modulo di avviso ottico e/o acustico (non illustrato).

Al dispositivo 1 è permesso l'upgrade dei software necessari per il suo buon funzionamento e la possibilità di controllare remotamente le funzioni della curva

di riferimento via Internet Protocol e connessione di rete.

Il dispositivo 1 è in grado di gestire vari tipi di segnali audio digitali, quali per esempio: audio contenuto in file audiovisivi memorizzati su supporto informatico nei più comuni formati/container (.mkv, .avi, .mov, e così via); segnali audio analogici prelevati, ad esempio, attraverso il "line out" di schede audio; segnali audio digitali prelevati tramite porta HDMI e/o BNC; segnali audio digitali prelevati tramite porta ottica (ADAT) e/o coassiale (SPDIF); ingresso microfonico/linea.

Il dispositivo 1 garantisce un ritardo massimo totale di attraversamento ed elaborazione dei segnali audio non percepibile acusticamente dall'utente, ovvero la capacità di ritardare il segnale video dell'esatta latenza di processamento del segnale audio, in modo da mantenere il sincronismo audio-video invariato tra ingresso e uscita al dispositivo 1. Inoltre, il dispositivo 1 è in grado di riprodurre agevolmente contenuti audiovisivi fino alla risoluzione HD.

Il dispositivo 1 è in grado di riprodurre segnali audio di tipo mono, stereo e i due canali frontali di un sistema multi-channel 5.1 oppure 7.1.

Il dispositivo 1 può funzionare in modalità stand-alone oppure essere collocato in una catena di trasmissione di un broadcaster.

Con riferimento alla Figura 6, viene illustrato uno schema a blocchi di componenti del processore 5 di segnale audio digitale che sono rilevanti per l'applicazione del metodo secondo l'invenzione, ossia: un modulo 20 per l'analisi multibanda di un segnale audio digitale; i mezzi di memoria 7 di tipo non volatile che contengono valori tabellari della curva di riferimento 30, ossia una funzione di trasferimento; un modulo di rilevamento 22 del livello di un segnale audio; un modulo comparatore di soglia 23; un modulo 24 di amplificazione/attenuazione di un segnale audio; e un modulo 26 di limitazione della loudness del segnale audio digitale.

Con riferimento alla Figura 7, viene illustrato un diagramma di flusso 100 di un metodo secondo la presente invenzione per migliorare la percezione della qualità di un segnale audio digitale.

Il primo segnale audio digitale 10 viene trattato all'interno del processore 5 di

segnale audio digitale secondo i seguenti passi di metodo:

- a) al passo 102 viene effettuata un'analisi multibanda del primo segnale audio digitale 10 da parte del modulo 20 per l'analisi multibanda di un segnale audio digitale e un rilevamento del livello del segnale audio per ciascuna banda di frequenza da parte del modulo di rilevamento 22 del livello di un segnale audio; il primo segnale audio digitale 10 viene suddiviso in una pluralità di bande di frequenza di un terzo di ottava e vengono considerate in particolare le seguenti frequenze di centro banda per l'equalizzazione: 100 Hz; 160 Hz; 250 Hz; 400 Hz; 630 Hz; 1000 Hz; 1600 Hz; 2500 Hz; 4000 Hz; 6300 Hz; 10000 Hz.
- b) al passo 104 per ciascuna banda di frequenza del primo segnale audio digitale 10, il relativo spettro viene confrontato con la curva di riferimento 30, o funzione di trasferimento, memorizzata nei mezzi di memoria 7 di tipo non volatile;
- c) al passo 106 viene verificato tramite un modulo comparatore di soglia 23 se il livello del primo segnale audio digitale 10 in una determinata banda di frequenza supera un determinato livello di soglia, o gate threshold; in caso affermativo, detto segnale audio digitale 10 viene amplificato o attenuato dinamicamente secondo il successivo passo d); in caso negativo, il livello di detto segnale in quella determinata banda di frequenza non viene modificato e il passo d) non viene applicato;
- d) al passo 108 il confronto tra il segnale di una determinata banda di frequenza e la curva di riferimento determina la sua correzione (amplificazione o attenuazione) tramite il modulo di amplificazione/attenuazione 24;
- e) al passo 110 il primo segnale audio digitale 10 viene amplificato o attenuato tramite il modulo di amplificazione/attenuazione 24 in modo che la sua loudness sia normalizzata ad un livello target predefinito, come ad esempio di  $-23.0 \pm 0,5$  LUFS (Loudness Units relative to Full Scale) indicato dalla raccomandazione EBU R128.

Eseguendo i passi del metodo sopra descritto, il segnale audio viene trattato dal processore 5 di segnale audio digitale in modo da presentare una curva a livello

sostanzialmente costante (curva piatta) nelle varie bande di frequenza.

Con riferimento alla Figura 8, viene illustrato il grafico della curva di riferimento 30, o funzione di trasferimento, implementata nei mezzi di memoria 7 di tipo non volatile del processore 5 di segnale audio digitale ed oggetto della presente invenzione.

Come visibile nella Figura 9, che replica la Figura 8, in corrispondenza di alcuni valori di frequenza rappresentati sull'asse delle ascisse, e più precisamente in corrispondenza dei valori di frequenza di centro banda di un terzo di ottava pari a 100 Hz, 160 Hz, 250 Hz, 400 Hz, 630 Hz, 1 kHz, 1,6 kHz, 2,5 kHz, 4 kHz, 6,3 kHz e 10 kHz sono rappresentati rispettivi cerchi 13 che indicano l'amplificazione o l'attenuazione massime che possono essere applicate al primo segnale audio digitale 10 ad una determinata banda di frequenze. Infatti, l'equalizzazione operata dal processore 5 di segnale audio digitale è dinamica e varia a seconda del contenuto frequenziale del segnale audio trattato in modo da mantenere sostanzialmente costante l'energia del segnale in ciascuna banda di frequenza.

La curva di riferimento rappresentata nelle Figure 8 e 9 amplifica maggiormente le alte frequenze, con un'enfasi massima a 4 kHz.

In Figura 10 è rappresentata una tabella che indica per ciascuna delle frequenze puntuali sopraindicate la massima attenuazione/amplificazione applicabile dal processore 5 di segnale audio digitale in una determinata banda di frequenza.

Più in particolare, la colonna "Centro [dB]" indica la posizione del centro del cerchio 13 in dB ad una determinata frequenza di centro banda; la colonna "Guadagno max [dB]" indica il massimo guadagno applicabile dal processore 5 di segnale audio digitale a quella determinata frequenza, ossia il diametro del cerchio 13; la colonna "Min [dB]" indica il livello più basso del cerchio 13; la colonna "Max [dB]" indica il livello più alto del cerchio 13.

In Figura 11 viene rappresentato un ingrandimento della Figura 9 in corrispondenza dell'intervallo di frequenze 4-16 kHz. Il cerchio 13 è centrato ad un valore di +2,7 dB (colonna "Centro [dB]" di Figura 10) e la funzione di trasferimento può essere attenuata/amplificata di 6 dB (colonna "Guadagno max

[dB]" di Figura 9) potendo quindi raggiungere valori compresi tra -3,3 dB e 8,7 dB in accordo alla tabella di Figura 10.

Analogamente, alla frequenza di 10 kHz, il centro del cerchio è posizionato a -4,1 dB e la funzione di trasferimento può raggiungere valori compresi tra -3,75 dB e -4,45 dB.

In sostanza, la curva di riferimento 30 implementata nel processore 5 di segnale digitale è una curva continua passante per i valori indicati nella colonna "Centro [dB]" di Figura 9 in corrispondenza dei rispettivi valori di frequenza indicati nella colonna "Frequenza [Hz]" di Figura 9, in cui sono possibili escursioni dovute ad una attenuazione o una amplificazione riportate nella colonna centrale "Guadagno max [dB]" di Figura 9. Nelle frequenze di centro banda di terzo di ottava intermedie tra i valori riportati nella colonna "Frequenza [Hz]" i valori della curva di riferimento 30 sono ottenuti tramite interpolazione con funzione spline cubica.

Tale attenuazione o amplificazione è ottenuta facendo in modo che l'energia del segnale rimanga sostanzialmente costante (curva piatta) in ciascuna banda di frequenza e che la loudness del segnale audio rispetti non ecceda il valore target predefinito.

Nelle Figure 12, 13 e 14 sono illustrati tre esempi di applicazione della funzione di trasferimento relativi rispettivamente ai tre generi "Parlato", "Musica e Canto" e "Sport".

In Figura 12 è rappresentato lo spettro del segnale in uscita (in linea continua) sovrapposto a quello del segnale in ingresso (linea tratteggiata) nel caso di una traccia audio del genere "Parlato". Dal confronto tra i due spettri risulta evidente un guadagno maggiore nelle bande centrate a 2,5 kHz e 4 kHz.

In Figura 13 è rappresentato lo spettro del segnale in uscita (in linea continua) sovrapposto a quello del segnale in ingresso (linea tratteggiata) nel caso di una traccia audio del genere "Musica e Canto". Dal confronto tra i due spettri il guadagno nelle bande tra 2 kHz e 4 kHz risulta meno evidente perché il segnale in ingresso presenta già alti guadagni.

In Figura 14 è rappresentato lo spettro del segnale in uscita (in linea continua)

sovrapposto a quello del segnale in ingresso (linea tratteggiata) nel caso di una traccia audio del genere "Sport". Lo spettro del segnale in uscita risulta molto simile a quello del genere "Parlato", essendo tuttavia presente il rumore di campo che si sovrappone al parlato.

Al fine di valutare il miglioramento oggettivo della qualità dell'audio percepita applicando il metodo oggetto della presente invenzione, la Richiedente ha sviluppato un sistema di valutazione in grado di effettuare misure oggettive per il calcolo dell'Objective Index (OI), ossia l'indice di valutazione del miglioramento della qualità audio con l'applicazione della curva di riferimento di Figura 8, correlato ad una scala di gradimento soggettiva. Il metodo del test soggettivo scelto fa riferimento alla norma ITU-R BS.1116-1 e prende il nome di "Double-blind triple stimulus with hidden reference".

Il Subjective Difference Grade (SDG) stimato a partire dal calcolo dell'OI presenta errori di stima variabili tra l'11% e il 13%.

I risultati dell'SDG evidenziano un miglioramento del 25,3% in media su tre televisori di marche Samsung, LG e Sony del tipo a schermo piatto e tre generi diversi ("Parlato", "Musica e Canto", e "Sport"). In particolare, sono stati raggiunti risultati pari anche al 30% di miglioramento di qualità percepita del segnale audio.

Alla luce della presente descrizione sono quindi evidenti i vantaggi della presente invenzione.

Il metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi, in particolare del tipo a schermo piatto, e relativo dispositivo, migliorano vantaggiosamente la qualità audio, in particolare del parlato, percepita da un utente di tale ricevitore e ciò indifferentemente dalla tipologia di contenuto veicolata dal segnale audio o dal modo in cui il segnale televisivo è emesso dal broadcaster, in particolare via televisione digitale terrestre, satellite o Internet.

Numerose sono le varianti possibili al metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi, in particolare del tipo a schermo piatto, e relativo dispositivo descritti come

esempio, senza per questo uscire dai principi di novità insiti nell'idea inventiva, così come è chiaro che nella sua attuazione pratica le forme dei dettagli illustrati potranno essere diverse, e gli stessi potranno essere sostituiti con degli elementi tecnicamente equivalenti.

Ad esempio, il metodo oggetto della presente invenzione può essere implementato in un ricevitore di segnali televisivi del tipo a schermo piatto, in particolare un televisore, comprendente il processore 5 di segnale audio digitale. Dunque, è facilmente comprensibile che la presente invenzione non è limitata ad un metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi, in particolare del tipo a schermo piatto, e relativo dispositivo descritti come esempio, ma è passibile di varie modificazioni, perfezionamenti, sostituzioni di parti ed elementi equivalenti senza però allontanarsi dall'idea dell'invenzione, così come è precisato meglio nelle seguenti rivendicazioni.

\* \* \* \* \*

## RIVENDICAZIONI

1. Metodo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi, in particolare del tipo a schermo piatto, comprendente i seguenti passi realizzati in un processore (5) di segnale audio digitale:

a) effettuare un'analisi multibanda di un segnale audio digitale (10) in ingresso a detto processore (5) da parte di un modulo (20) per l'analisi multibanda suddividendo detto segnale audio digitale (10) in una pluralità di bande di frequenza;

b) rilevare un livello di segnale audio digitale (10) per ciascuna banda di frequenza da parte di un modulo di rilevamento (22);

c) per ciascuna banda di frequenza di detto segnale audio digitale (10);

c1) confrontare il relativo spettro con una curva di riferimento (30), memorizzata in mezzi di memoria (7) di tipo non volatile di detto processore (5);

c2) verificare per mezzo di un modulo comparatore di soglia (23) se il livello del segnale audio digitale (10) in una determinata banda di frequenza supera un determinato livello di soglia di detta curva di riferimento (30) e, in caso affermativo, amplificare/attenuare dinamicamente il livello di detto segnale audio digitale (10) in detta banda di frequenza per mezzo di un modulo di amplificazione/attenuazione (24);

d) amplificare/attenuare per mezzo di detto modulo di amplificazione/attenuazione (24) il livello di detto segnale audio digitale (10) in modo che la sua loudness sia normalizzata ad un livello target predefinito, detto metodo essendo caratterizzato dal fatto che

in detto step b) detto segnale audio digitale (10) viene suddiviso in bande di un terzo di ottava aventi rispettive frequenze di centro banda centrate nei valori di frequenza: 100 Hz; 160 Hz; 250 Hz; 400 Hz; 630 Hz; 1000 Hz; 1600 Hz; 2500 Hz; 4000 Hz; 6300 Hz; 10000 Hz;

dal fatto che detto livello di soglia di detta curva di riferimento (30) assume i

seguenti valori: a 100 Hz, -2,10 dB; a 160 Hz, -7,50 dB; a 250 Hz, -0,15 dB; a 400 Hz, -2,50 dB; a 630 Hz, -2,65 dB; a 1000 Hz, -1,80 dB; a 1600 Hz, -3,55 dB; a 2500 Hz, -1,55 dB; a 4000 Hz, -3,30 dB; a 6300 Hz, -0,85 dB; a 10000 Hz, -4,45 dB; e dal fatto che il livello del segnale audio digitale (10) è amplificato/attenuato nei seguenti intervalli di valori: a 100 Hz, -0,10 dB/-2,10 dB; a 160 Hz, -7,50/-5,90 dB; a 250 Hz, -0,15 dB/1,35 dB; a 400 Hz, -2,50 dB/-0,30dB; a 630 Hz, -2,65 dB/1,05 dB; a 1000 Hz, -1,80 dB/4dB ; a 1600 Hz, -3,55 dB/2,55 dB; a 2500 Hz, 1,55 dB/7,05 dB; a 4000 Hz, -3,30 dB/8,70 dB; a 6300 Hz, -0,85 dB/0,25 dB; a 10000 Hz, -4,45 dB/-3,75 dB.

2. Metodo secondo la rivendicazione 1, in cui detto livello target predefinito è  $-23.0 \pm 0,5$  LUFS.

3. Metodo secondo la rivendicazione 2, in cui nei valori di frequenza compresi tra detti valori di centro banda, i valori di detta curva di riferimento (30) sono ottenuti tramite interpolazione con funzione spline cubica.

4. Dispositivo per migliorare la percezione della qualità di un segnale audio digitale emesso da un ricevitore di segnali televisivi, in particolare del tipo a schermo piatto, comprendente:

- mezzi di ingresso (4) di un segnale audio/video digitale (3);
- mezzi (9) per estrarre un segnale audio digitale (10) da detto segnale audio/video digitale (3);
- un blocco estrattore/ritardatore (13) per estrarre un segnale video (6) da detto segnale audio/video digitale (3) e ritardarlo;
- un processore (5) di detto segnale audio digitale (10) comprendente mezzi di memoria (7) di tipo non volatile;
- mezzi (12) per ricombinare un secondo segnale audio digitale (11) in uscita da detto processore (5), in particolare un de-embedder, con detto segnale video (6), in cui detto processore (5) comprende:
  - un modulo (20) per effettuare l'analisi multibanda di detto segnale audio digitale (10) suddividendo detto segnale audio digitale (10) in una pluralità di bande di frequenza;
  - un modulo di rilevamento (22) per rilevare un livello di detto segnale audio

digitale (10) per ciascuna banda di frequenza;

- mezzi per confrontare uno spettro di ciascuna banda di frequenza di detto segnale audio digitale (10) con una curva di riferimento (30), memorizzata in detti mezzi di memoria (7) di tipo non volatile;

- un modulo comparatore di soglia (23) atto a verificare se il livello di detto segnale audio digitale (10) in ciascuna banda di frequenza supera un determinato livello di soglia di detta curva di riferimento (30);

- un modulo di amplificazione/attenuazione (24) atto ad amplificare/attenuare dinamicamente il livello di detto segnale audio digitale (10) in detta banda di frequenza in modo che la loudness del segnale audio digitale sia normalizzata ad un livello target predefinito,

detto dispositivo essendo caratterizzato dal fatto che detto modulo (20) per effettuare l'analisi multibanda di detto segnale audio digitale (10) è atto a suddividere detto segnale audio digitale (10) in bande di un terzo di ottava aventi rispettive frequenze di centro banda centrate nei valori di frequenza: 100 Hz; 160 Hz; 250 Hz; 400 Hz; 630 Hz; 1000 Hz; 1600 Hz; 2500 Hz; 4000 Hz; 6300 Hz; 10000 Hz;

dal fatto che detto livello di soglia di detta curva di riferimento (30) assume i seguenti valori: a 100 Hz, -2,10 dB; a 160 Hz, -7,50 dB; a 250 Hz, -0,15 dB; a 400 Hz, -2,50 dB; a 630 Hz, -2,65 dB; a 1000 Hz, -1,80 dB; a 1600 Hz, -3,55 dB; a 2500 Hz, -1,55 dB; a 4000 Hz, -3,30 dB; a 6300 Hz, -0,85 dB; a 10000 Hz, -4,45 dB; e dal fatto che detto modulo di amplificazione/attenuazione (24) è atto a amplificare/attenuare il livello del segnale audio digitale (10) nei seguenti intervalli di valori: a 100 Hz, -0,10 dB/-2,10 dB; a 160 Hz, -7,50/-5,90 dB; a 250 Hz, -0,15 dB/1,35 dB; a 400 Hz, -2,50 dB/-0,30dB; a 630 Hz, -2,65 dB/1,05 dB; a 1000 Hz, -1,80 dB/4dB ; a 1600 Hz, -3,55 dB/2,55 dB; a 2500 Hz, 1,55 dB/7,05 dB; a 4000 Hz, -3,30 dB/8,70 dB; a 6300 Hz, -0,85 dB/0,25 dB; a 10000 Hz, -4,45 dB/-3,75 dB.

5. Dispositivo secondo la rivendicazione 4, in cui detto livello target predefinito è  $-23.0 \pm 0,5$  LUFS.

6. Dispositivo secondo la rivendicazione 4 o 5, in cui detto dispositivo comprende

un modulo di by-pass automatico che si attiva in caso di avaria di detto processore (5) di segnale audio digitale oppure dell'intero dispositivo (1).

7. Dispositivo secondo la rivendicazione 6, in cui detto dispositivo comprende mezzi di segnalazione, in particolare un modulo di avviso ottico e/o acustico, atto a segnalare detta avaria.

8. Dispositivo secondo una delle rivendicazioni da 4 a 7, in cui detto dispositivo comprende una connessione di rete per aggiornare il software necessario al funzionamento di detto dispositivo e controllare remotamente le funzioni di detta curva di riferimento (30).

9. Ricevitore di segnali televisivi del tipo a schermo piatto, in particolare un televisore, comprendente un processore (5) di segnale audio digitale atto ad implementare il metodo secondo una o più delle rivendicazioni da 1 a 3.

Sesso del parlatore	Frequenza centrale delle bande di ottava (Hz)							
	125	250	500	1000	2000	4000	8000	
Maschio	$\alpha$	0,085	0,127	0,230	0,233	0,309	0,224	0,173
Femmina	$\alpha$	-	0,117	0,223	0,216	0,328	0,250	0,194

Fig. 1

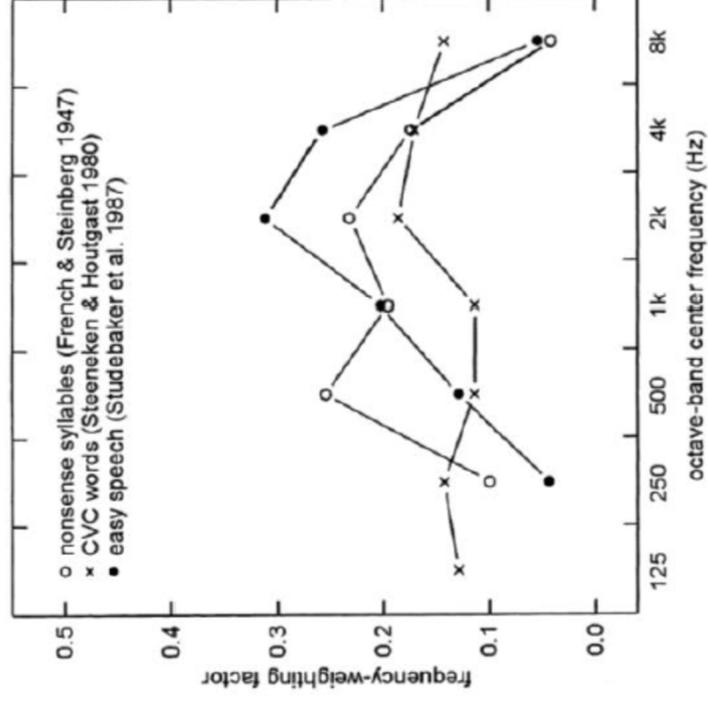


Fig. 2

**Octave band importance functions for various speech tests.**

Band No.	Midband Freq, Hz	NNS <sup>a</sup>	CID-22 <sup>b</sup>	NU6 <sup>c</sup>	DRT <sup>d</sup>	Short Passages <sup>e</sup>	SPIN <sup>f</sup>
1	250	0.0437	0.1549	0.0853	0.0960	0.1004	0.0871
2	500	0.1294	0.1562	0.1912	0.2043	0.2551	0.1493
3	1000	0.2025	0.2165	0.2110	0.2343	0.1960	0.2206
4	2000	0.3117	0.2768	0.3090	0.2643	0.2322	0.3022
5	4000	0.2576	0.1488	0.1682	0.1501	0.1744	0.2102
6	8000	0.0551	0.0468	0.0353	0.0510	0.0419	0.0306

<sup>a</sup>NNS (various nonsense syllable tests where most of the English phonemes occur equally often), <sup>b</sup>CID-W22 (PB-words), <sup>c</sup>NU6 monosyllables, <sup>d</sup>DRT (Diagnostic Rhyme Test), <sup>e</sup>short passages of easy reading material, <sup>f</sup>SPIN monosyllables

**Fig. 3**

**Table B.2 — One-third octave band importance functions for various speech tests.**

Band No.	Midband Freq, Hz	NNS <sup>a</sup>	CID-22 <sup>b</sup>	NU6 <sup>c</sup>	DRT <sup>d</sup>	Short Passages <sup>e</sup>	SPIN <sup>f</sup>
1	160	0.0000	0.0365	0.0168	0.0000	0.0114	0.0000
2	200	0.0000	0.0279	0.0130	0.0240	0.0153	0.0255
3	250	0.0153	0.0405	0.0211	0.0330	0.0179	0.0256
4	315	0.0284	0.0500	0.0344	0.0390	0.0558	0.0360
5	400	0.0363	0.0530	0.0517	0.0571	0.0898	0.0362
6	500	0.0422	0.0518	0.0737	0.0691	0.0944	0.0514
7	630	0.0509	0.0514	0.0658	0.0781	0.0709	0.0616
8	800	0.0584	0.0575	0.0644	0.0751	0.0660	0.0770
9	1000	0.0667	0.0717	0.0664	0.0781	0.0628	0.0718
10	1250	0.0774	0.0873	0.0802	0.0811	0.0672	0.0718
11	1600	0.0893	0.0902	0.0987	0.0961	0.0747	0.1075
12	2000	0.1104	0.0938	0.1171	0.0901	0.0755	0.0921
13	2500	0.1120	0.0928	0.0932	0.0781	0.0820	0.1026
14	3150	0.0981	0.0678	0.0783	0.0691	0.0808	0.0922
15	4000	0.0867	0.0498	0.0562	0.0480	0.0483	0.0719
16	5000	0.0728	0.0312	0.0337	0.0330	0.0453	0.0461
17	6300	0.0551	0.0215	0.0177	0.0270	0.0274	0.0306
18	8000	0.0000	0.0253	0.0176	0.0240	0.0145	0.0000

<sup>a</sup>NNS (various nonsense syllable tests where most of the English phonemes occur equally often), <sup>b</sup>CID-W22 (PB-words), <sup>c</sup>NU6 monosyllables, <sup>d</sup>DRT (Diagnostic Rhyme Test), <sup>e</sup>short passages of easy reading material, <sup>f</sup>SPIN monosyllables

**Fig. 4**

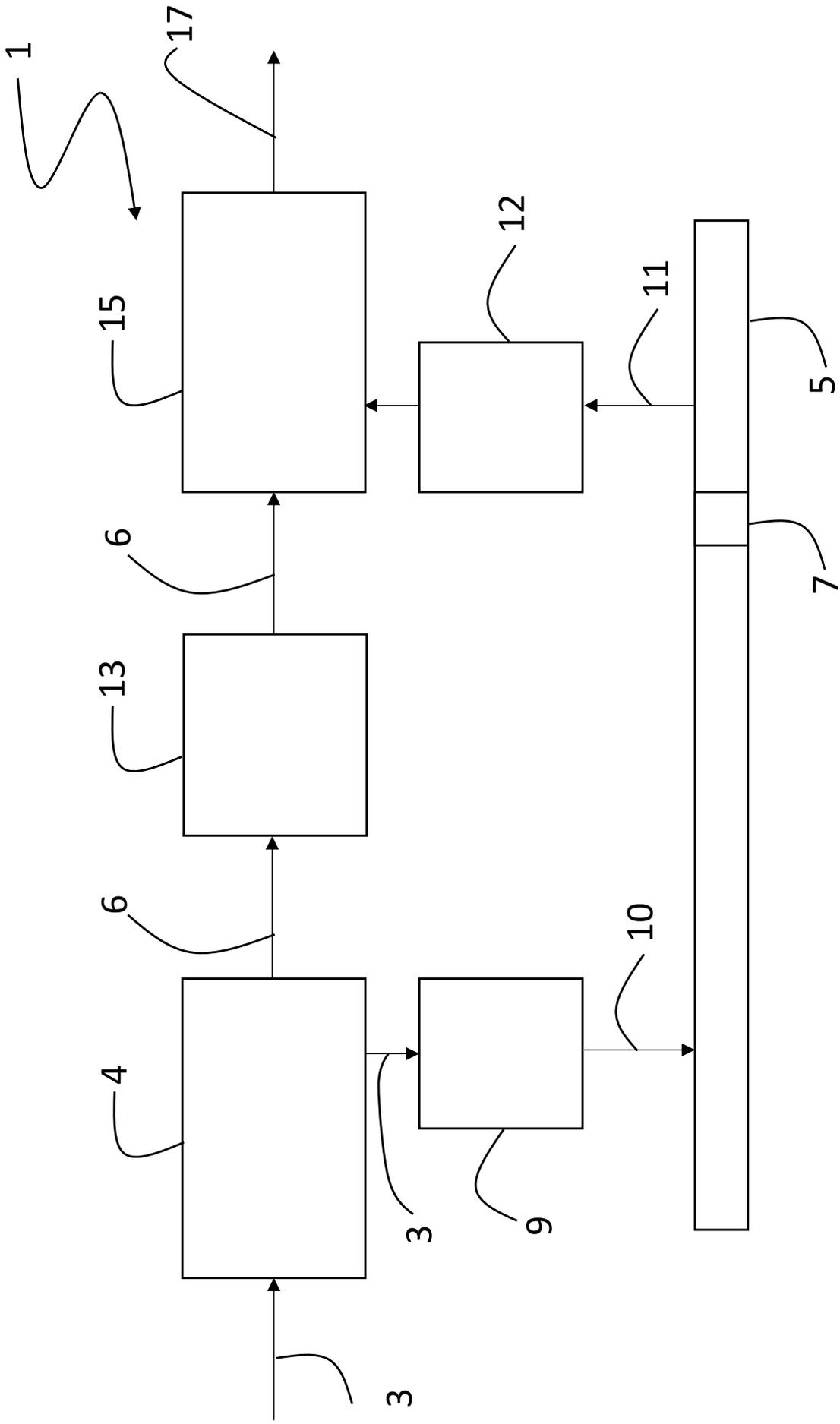


Fig. 5

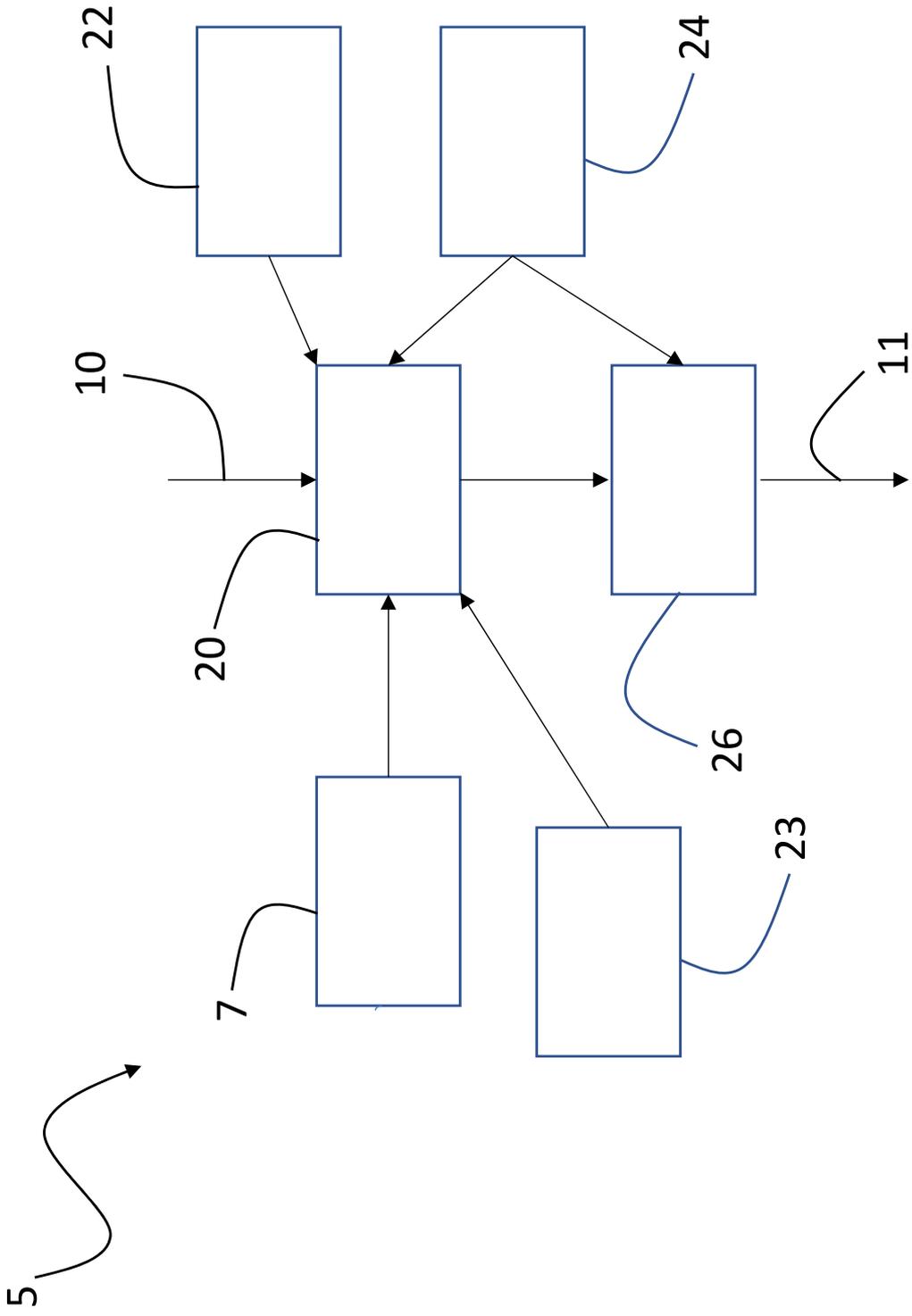


Fig. 6

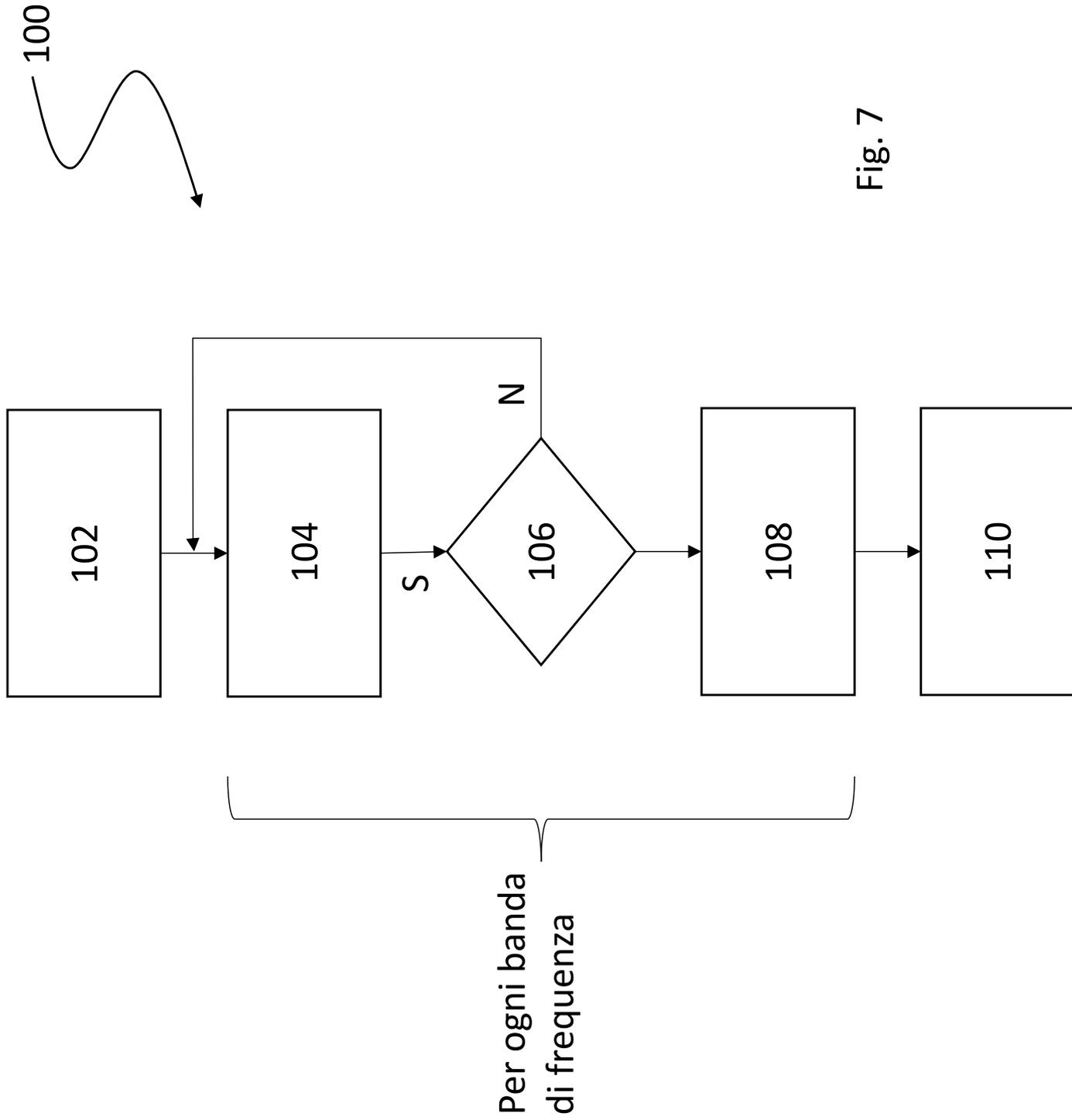


Fig. 7

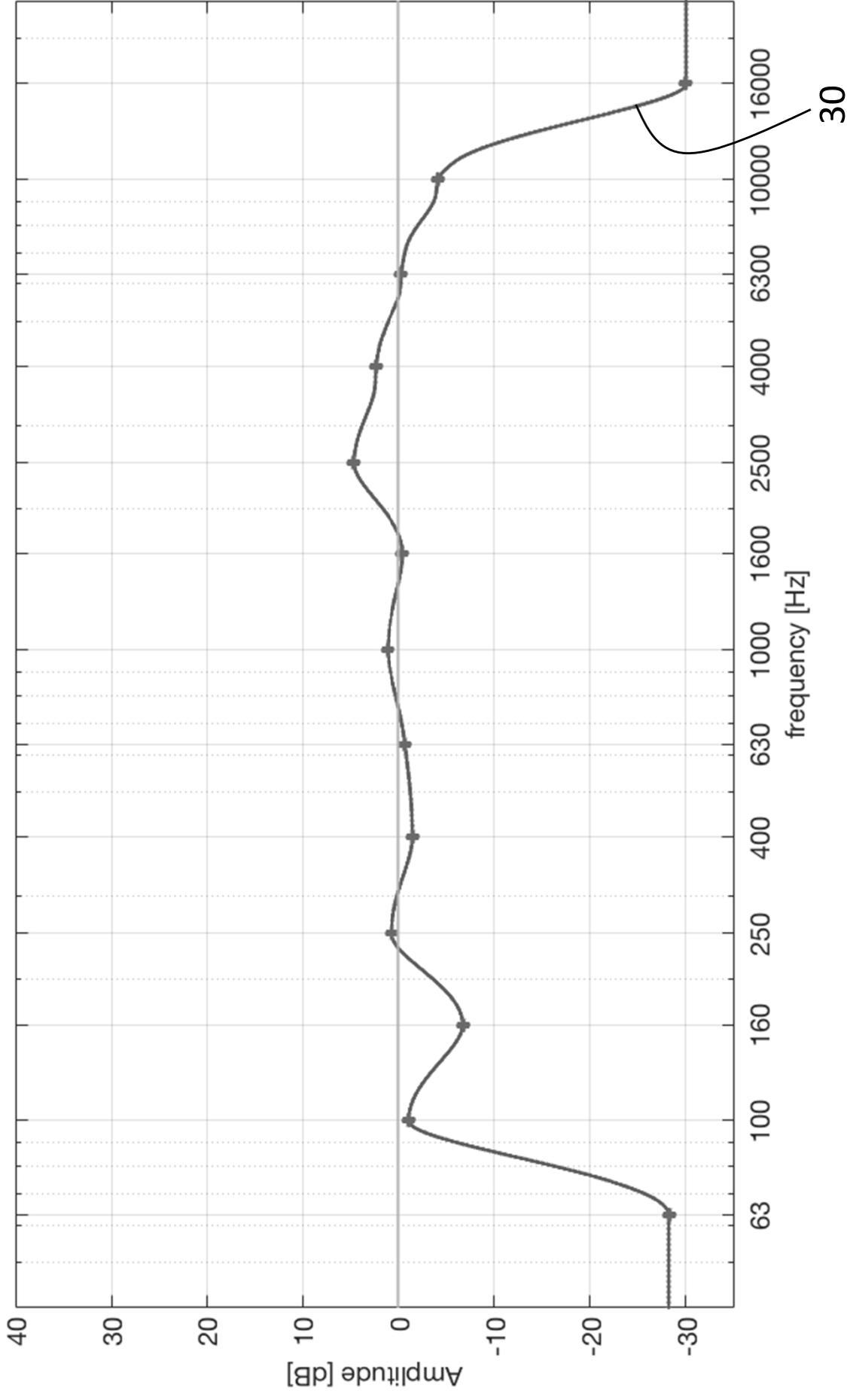


Fig. 8

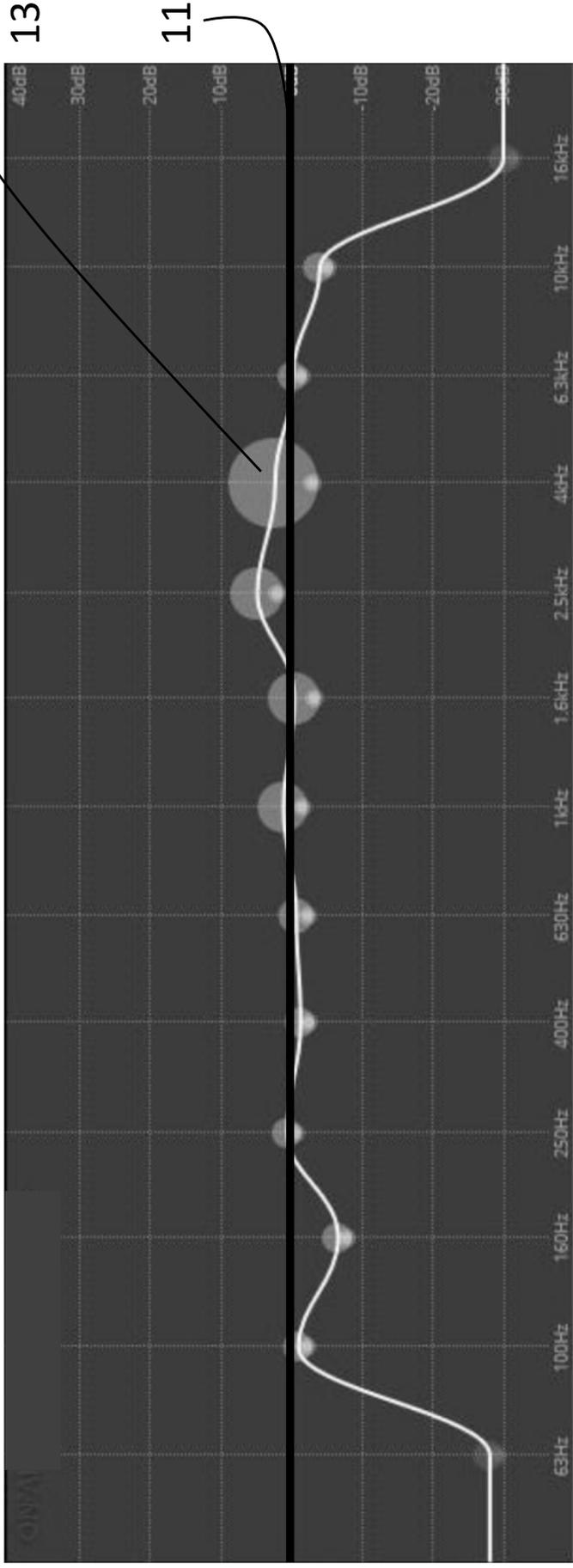


Fig. 9

Freq [Hz]	Min [dB]	Centro [dB]	Max [dB]	Guadagno max [dB]
100	-2,10	-1,10	-0,10	2,0
160	-7,50	-6,70	-5,90	1,6
250	-0,15	0,60	1,35	1,5
400	-2,50	-1,40	-0,30	2,2
630	-2,65	-0,80	1,05	3,7
1000	-1,80	1,10	4,00	5,8
1600	-3,55	-0,50	2,55	6,1
2500	1,55	4,30	7,05	5,5
4000	-3,30	2,70	8,70	12,0
6300	-0,85	-0,30	0,25	1,1
10000	-4,45	-4,10	-3,75	0,7

Fig. 10

13

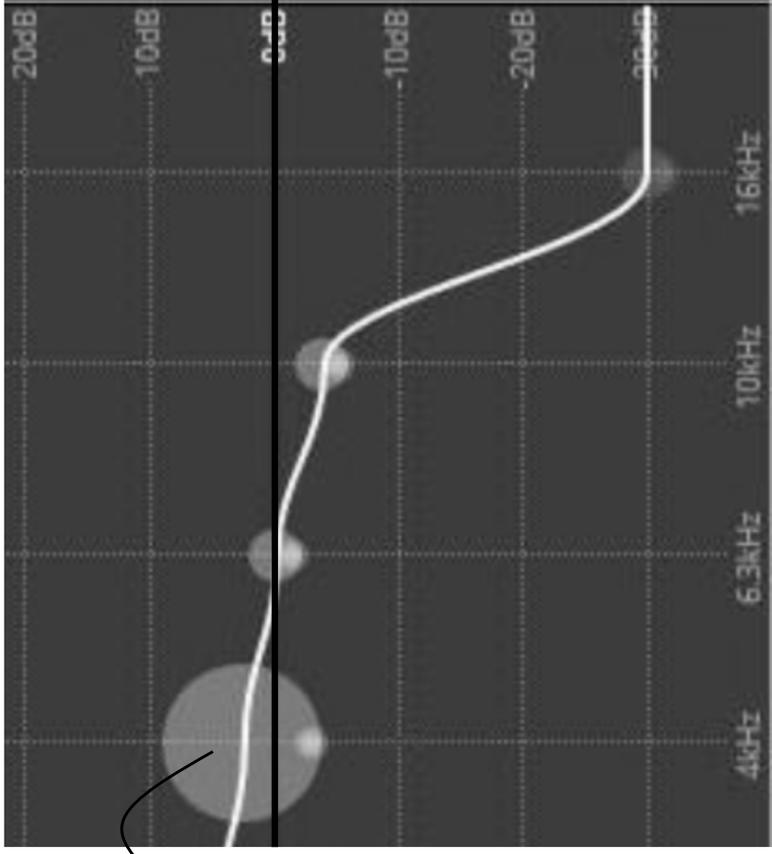


Fig. 11

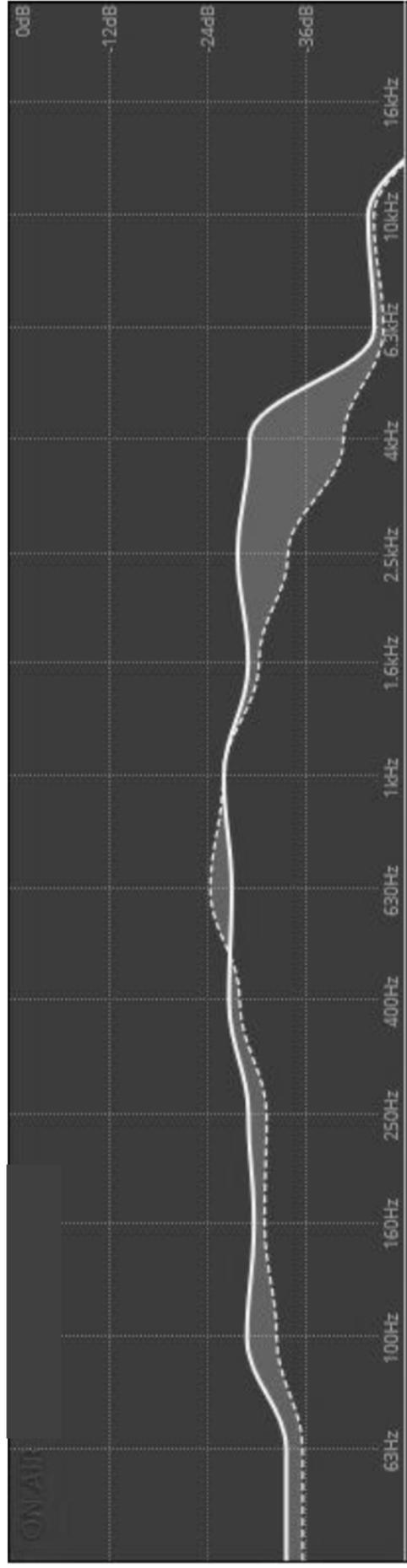


Fig. 12

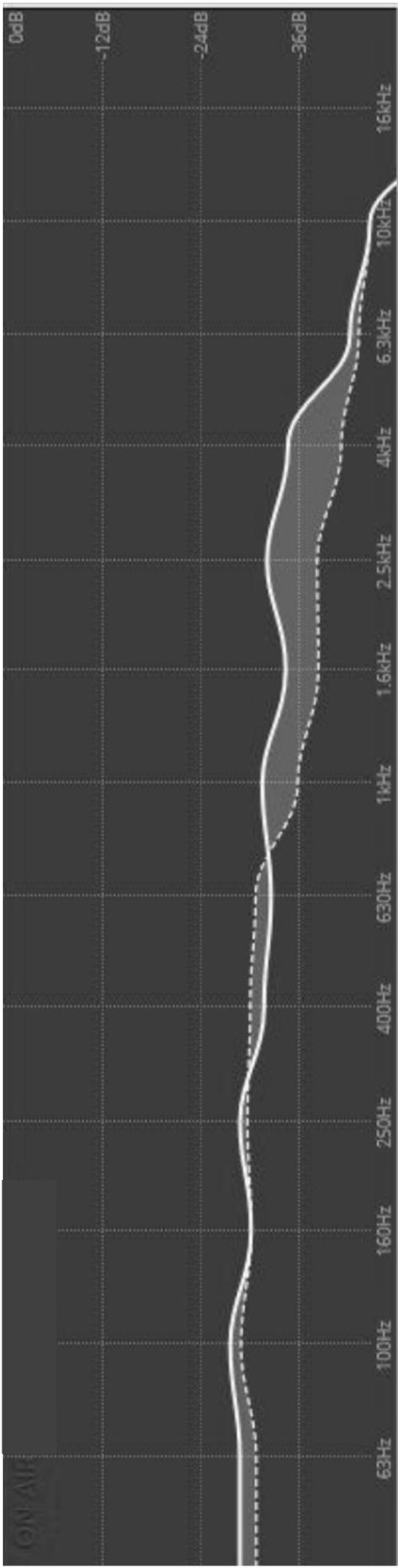


Fig. 13

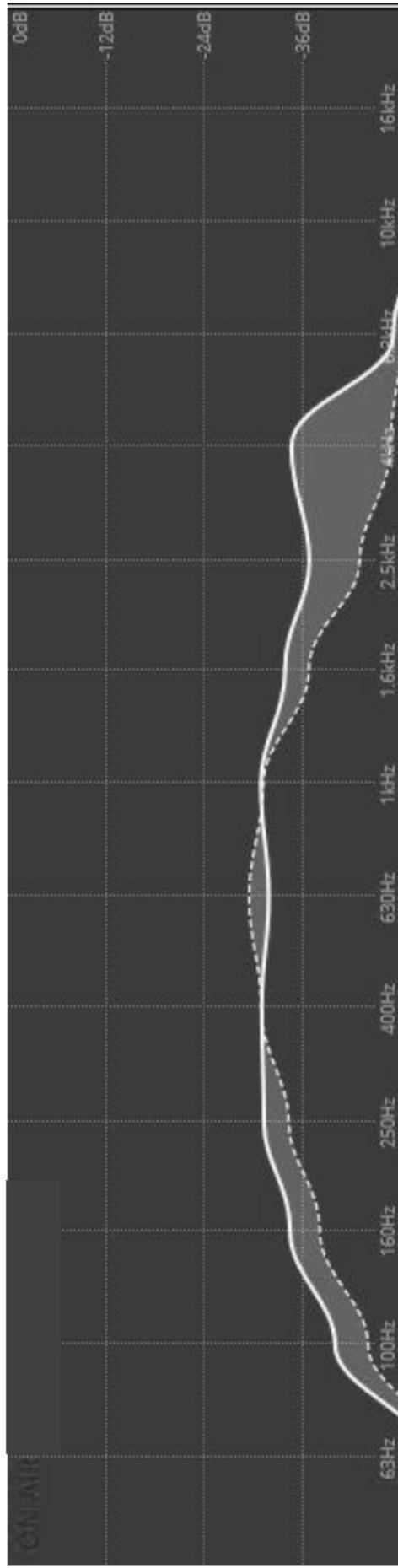


Fig. 14