



MINISTERO DELLO SVILUPPO ECONOMICO
DIREZIONE GENERALE PER LA TUTELA DELLA PROPRIETÀ INDUSTRIALE
UFFICIO ITALIANO BREVETTI E MARCHI

UIBM

DOMANDA NUMERO	101994900393024
Data Deposito	29/09/1994
Data Pubblicazione	29/03/1996

Sezione	Classe	Sottoclasse	Gruppo	Sottogruppo
G	10	L		

Titolo

PROCEDIMENTO DI SINTESI VOCALE MEDIANTE CONCATENAZIONE E PARZIALE SOVRAPPOSIZIONE DI FORME D'ONDA.

Descrizione dell'invenzione avente per titolo:

"PROCEDIMENTO DI SINTESI VOCALE MEDIANTE CONCATENAZIONE E PARZIALE
SOVRAPPOSIZIONE DI FORME D'ONDA"

a nome CSELT - Centro Studi e Laboratori Telecomunicazioni S.p.A., Via G. Reiss
Romoli 274, 10148 Torino, nazionalità italiana.

Inventori : Enzo FOTI

Luciano NEBBIA

Stefano SANDRI

Domanda N.

TO 94A00075G

Depositata il 29.8.1984

=====

La presente invenzione si riferisce alla sintesi della voce, e in particolare ha per oggetto un procedimento e un dispositivo di sintesi basato sulla concatenazione di forme d'onda relative ad unità elementari di voce. Preferibilmente, ma non esclusivamente, il procedimento e il dispositivo trovano applicazione nella sintesi da testo scritto.

In queste applicazioni, un testo da trasformare in segnale vocale viene dapprima convertito in una rappresentazione fonetico-prosodica, che indica la sequenza dei fonemi corrispondenti e le caratteristiche prosodiche (durata, intensità e periodo fondamentale) ad essi associate. Tale rappresentazione è poi convertita in un segnale vocale sintetico numerico a partire da un vocabolario di dette unità elementari, che nel caso più frequente sono costituite da difoni (elementi di voce che si estendono dalla parte stazionaria di un fonema alla parte stazionaria del fonema successivo, includendo la transizione tra di essi). Nel caso della lingua italiana, un vocabolario di circa un migliaio di difoni assicura la copertura fonetica permettendo di sintetizzare tutti i suoni ammissibili per l'italiano.

Nei sistemi di sintesi da testo scritto, per la generazione del segnale vocale si possono utilizzare metodi basati sulla concatenazione, nel dominio del tempo, delle forme d'onda che rappresentano le varie unità elementari: questi metodi sono molto flessibili e garantiscono una buona qualità della voce sintetica.

Un esempio è descritto da E. Moulines e F. Charpentier nell'articolo "Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones", Speech Communication, Vol. 9, N. 5/6, Dicembre 1990, pagg. 453-467. Questo metodo si basa sulla tecnica nota come PSOLA (Pitch-Synchronous OverLap and Add : sovrapposizione e somma sincrone con il periodo fondamentale), per applicare la prosodia imposta dalle regole di sintesi e concatenare le forme d'onda delle unità elementari. La tecnica PSOLA effettua, almeno per i tratti vocalizzati del segnale originale, un'analisi mediante l'applicazione di una finestrazione sincrona con il periodo fondamentale, in particolare utilizzando finestre di Hanning di durata prossima al doppio del periodo fondamentale, generando una sequenza di segnali a breve termine parzialmente sovrapposti; in fase di sintesi i segnali risultanti dalla finestrazione vengono traslati temporalmente in modo sincrono con il periodo fondamentale imposto dalle regole prosodiche per la sintesi; infine la generazione del segnale sintetico è effettuata per sovrapposizione e somma dei segnali traslati. Per ridurre la complessità dei calcoli, il secondo passo può essere fatto direttamente nel dominio del tempo.

La finestrazione completa dei singoli intervalli del segnale originale richiede un carico di calcolo relativamente elevato e inoltre, costituendo un'alterazione del segnale originale che si estende sull'intero intervallo, comporta una diminuzione di naturalezza sul segnale sintetico.

Secondo l'invenzione si fornisce un procedimento di sintesi mediante il quale si lascia invariata quella parte di ogni intervallo del segnale originale che contiene le

informazioni fondamentali, e si altera solo la parte restante dell'intervallo: in questo modo non solo si riduce il tempo di elaborazione, ma si migliora anche la naturalezza del segnale sintetico, in quanto la parte principale dell'intervallo è la riproduzione esatta del segnale originale.

L'invenzione fornisce pertanto un procedimento per la sintesi di segnali vocali mediante concatenazione nel tempo di forme d'onda che rappresentano unità elementari di segnale vocale, in cui: si suddividono almeno le forme d'onda associate a suoni vocalizzati in una pluralità di intervalli, corrispondenti alle risposte del condotto vocale a una serie di impulsi di eccitazione delle corde vocali sincroni con la frequenza fondamentale del segnale; si sottopone a pesatura la forma d'onda compresa in ciascun intervallo; si sostituiscono i segnali risultanti dalla pesatura con una loro replica sfalsata nel tempo in misura dipendente da un'informazione prosodica; e si effettua la sintesi mediante sovrapposizione e somma dei segnali sfalsati; e in cui:

- si suddivide un intervallo corrente di segnale originale che deve essere riprodotto in sintesi in una parte invariante, che è compresa tra l'inizio dell'intervallo e un bordo sinistro di analisi, rappresentato da un passaggio per lo zero del segnale vocale originale che soddisfa a condizioni prestabilite, e una parte variabile, che si estende tra il bordo sinistro di analisi e un bordo destro di analisi sostanzialmente coincidente con la fine dell'intervallo corrente, il bordo sinistro e il bordo destro di analisi essendo associati, nel segnale sintetizzato, rispettivamente a un bordo sinistro di sintesi e a un bordo destro di sintesi, dei quali il primo coincide con il bordo sinistro di analisi, con riferimento a un marcatore di inizio intervallo, e il secondo coincide sostanzialmente con la fine dell'intervallo nel segnale sintetizzato;
- si applica, sulla parte di forma d'onda a destra del bordo sinistro di analisi dell'intervallo corrente di segnale originale, una prima funzione di raccordo, che ha

- durata uguale a quella del tratto di forma d'onda sintetizzata compreso tra il bordo sinistro e il bordo destro di sintesi e ampiezza progressivamente decrescente e massima in corrispondenza del bordo sinistro di analisi;
- si applica, sulla parte di forma d'onda a sinistra dell'intervallo immediatamente successivo di segnale originale da riprodurre in sintesi, una seconda funzione di raccordo, che ha durata uguale a quella del tratto di forma d'onda sintetizzata compreso tra il bordo sinistro e il bordo destro di sintesi e ampiezza progressivamente crescente e massima in corrispondenza dell'inizio di detto intervallo successivo;
 - si costruisce ogni intervallo di segnale sintetizzato riproducendo immutata la forma d'onda compresa nella parte invariante dell'intervallo originale, a cui viene unita la forma d'onda ottenuta allineando nel tempo e sommando le due forme d'onda risultanti dall'applicazione della prima e seconda funzione di raccordo.

A maggior chiarimento si fa riferimento ai disegni allegati, che illustrano un esempio di realizzazione dell'invenzione dato a titolo non limitativo e in cui:

- la fig. 1 è lo schema generale delle operazioni di un sistema di sintesi da testo scritto per concatenazione di unità acustiche elementari;
- la fig. 2 è un diagramma del procedimento di sintesi per concatenazione di difoni e modifica dei parametri prosodici nel dominio temporale, secondo l'invenzione;
- la fig. 3 rappresenta la forma d'onda di un difono reale, con le marcature di confine fonetico e di difono e le marcature di periodicità;
- le figure 4, 5 e 6 sono grafici che rappresentano come si modificano i parametri prosodici di un segnale vocale naturale in alcuni casi particolari, secondo l'invenzione;
- le figure 7A, 7B, 8A, 8B, 9A, 9B, 10A e 10B sono alcuni esempi reali di

- applicazione del procedimento secondo l'invenzione per la modifica del periodo fondamentale su tratti del difono di fig. 3;
- le figure 11 - 18 sono diagrammi di flusso delle operazioni di determinazione del bordo sinistro di analisi e di sintesi

Prima di descrivere nei dettagli l'invenzione, si illustra brevemente la struttura di un sistema di sintesi vocale a partire da un testo scritto.

Come si vede in fig. 1, come prima fase il testo scritto viene fornito a uno stadio di trattamento linguistico TL che trasforma il testo scritto in una forma pronunciabile e con l'aggiunta di marcature linguistiche: trascrizione di sigle, numeri, ..., applicazione di regole di accentazione e classificazione grammaticale, accesso a informazioni lessicali contenute in un apposito vocabolario VL. Lo stadio successivo TF effettua la trascrizione da sequenza ortografica alla corrispondente stringa di simboli fonetici. Lo stadio di trattamento prosodico TP, sulla base di una serie di regole prosodiche RP, fornisce per ognuno dei fonemi in uscita da TF la durata e il periodo (e quindi la frequenza) fondamentale. Queste informazioni sono poi fornite allo stadio di pre-sintesi PS, che determina per ogni fonema la sequenza di segmenti acustici che lo compongono (accesso alla base dati dei difoni VD) e, per ogni segmento, quanti e quali intervalli, di durata uguale al periodo fondamentale, si devono utilizzare (nel caso di suoni vocalizzati) e i corrispondenti valori del periodo fondamentale da attribuire in sintesi, ottenuti per interpolazione dei valori assegnati ai confini di fonema; nel caso di suoni non vocalizzati, in cui non esistono caratteristiche di periodicità, gli intervalli hanno durata fissa. Queste informazioni sono infine utilizzate dal sintetizzatore vero e proprio SINT che opera le trasformazioni necessarie per la generazione del segnale sintetico.

La fig. 2 illustra in maggiore dettaglio il funzionamento dei moduli PS e SINT.

L'ingresso è costituito dall'identificativo del fonema corrente F_i , dalla sua durata D_i e dal valore del periodo fondamentale P_{i-1} all'inizio del fonema e P_i a fine fonema e dagli identificativi del fonema precedente F_{i-1} e seguente F_{i+1} . La prima operazione che viene effettuata è la decodifica del difono DF_{i-1} e del difono DF_i e la rilevazione dei marcatori di inizio e fine difono e di confine di fonema. Queste informazioni vengono prelevate direttamente dalla base dati dove sono memorizzati i difoni in forma d'onda e i relativi descrittori di confine, di decisione vocalizzato/non vocalizzato e di marcatura di periodicità. Il modulo successivo trasforma tali descrittori prendendo come riferimento il fonema. Sulla base di queste informazioni un modulo ritmico calcola il rapporto tra la durata D_i imposta dalla regola e la durata intrinseca del fonema (memorizzata nel vocabolario e data dalla somma delle due porzioni del fonema appartenente ai due difoni DF_{i-1} e DF_i). Quindi, tenendo conto della modifica della durata, calcola il numero degli intervalli da utilizzare in sintesi e determina per ciascuno di essi il valore del periodo fondamentale, mediante una legge di interpolazione tra il valore P_{i-1} e P_i . Il valore del periodo fondamentale è poi effettivamente utilizzato solo per i suoni vocalizzati, mentre per i suoni non vocalizzati, come detto sopra, si considerano intervalli di durata fissa.

Per la sintesi vera e propria, si procede in modo diverso a seconda che il suono sia vocalizzato o no.

Nel caso di suono non vocalizzato, il procedimento di sintesi consiste in un semplice scalamento temporale (allungamento o accorciamento) degli intervalli suddetti in base al rapporto tra la durata imposta dalle regole prosodiche e quella intrinseca. Nel caso di suono vocalizzato, si applica invece il procedimento oggetto dell'invenzione.

Il procedimento di sintesi secondo l'invenzione parte dalla considerazione che un

suono vocalizzato può essere considerato come sequenza di intervalli quasi periodici, ciascuno definito da un valore del periodo fondamentale p_a . Questo si vede chiaramente nella fig. 3 che riporta la forma d'onda del difono "à_m", i relativi marcatori che separano i singoli intervalli e, per ciascuno di questi, il valore del corrispondente periodo p_a espresso in Hz. La parte della fig. 3 compresa tra i due marcatori "v" corrisponde alla porzione destra del fonema "à"; la parte compresa tra il secondo marcatore "v" e il marcatore di fine difono "f" corrisponde alla porzione sinistra del fonema "m". Gli intervalli suddetti sono assimilabili alle risposte all'impulso di un filtro, stazionario per alcuni millisecondi e corrispondente al tratto vocale, che viene eccitato da una sequenza di impulsi disposti temporalmente in modo sincrono con la frequenza fondamentale della sorgente (frequenza di vibrazione delle corde vocali). L'operazione che deve compiere il modulo di sintesi, per ogni intervallo, è di ricevere il segnale originale con periodo fondamentale p_a (periodo di analisi) e di riprodurre in uscita un segnale modificato con il periodo p_s (periodo di sintesi) richiesto dalle regole prosodiche.

Le informazioni essenziali che caratterizzano ciascun intervallo vocale sono contenute nella zona di segnale che segue immediatamente l'impulso di eccitazione (parte principale della risposta), mentre la risposta stessa diventa sempre meno significativa via via che ci si allontana dalla posizione dell'impulso. Tenuto conto di ciò, nel procedimento di sintesi secondo l'invenzione si mantiene invariata il più possibile tale parte principale e l'allungamento o l'accorciamento del periodo richiesto dalle regole prosodiche vengono ottenuti agendo sulla parte residua.

A questo scopo si individuano quindi in ciascun intervallo una parte invariante e una parte variabile, e soltanto quest'ultima viene interessata ad operazioni di raccordo, sovrapposizione e somma. La parte invariante del segnale originale non è

costante, ma dipende per ogni intervallo dal rapporto tra p_s e p_a . Tale parte invariante è compresa tra il marcatore di inizio dell'intervallo e un cosiddetto bordo sinistro di analisi b_{sa} , che è costituito da uno dei passaggi per lo zero del segnale vocale originale, individuato con criteri che saranno descritti in seguito e che possono essere diversi a seconda che il periodo di sintesi sia maggiore oppure minore o uguale a quello di analisi. La parte variabile è delimitata dal bordo sinistro di analisi b_{sa} e da un cosiddetto bordo destro di analisi b_{da} , che coincide sostanzialmente con la fine dell'intervallo, in particolare con il campione precedente al marcatore di inizio dell'intervallo successivo.

Nel segnale sintetizzato, al bordo sinistro e al bordo destro di analisi b_{sa} , b_{da} sono associati un bordo sinistro e un bordo destro di sintesi b_{ss} , b_{ds} . Per un dato intervallo il bordo sinistro di sintesi ovviamente coincide con quello di analisi, con riferimento al marcatore d'inizio dell'intervallo, in quanto la parte di segnale che precede viene riprodotta inalterata nella sintesi. Il bordo destro di sintesi è definito dalla relazione:

$$b_{ds} = b_{da} + \Delta p \quad (1)$$

dove $\Delta p = p_s - p_a$ avrà valore positivo o negativo a seconda che in sintesi si abbia un allungamento o un accorciamento del periodo fondamentale.

La parte variabile dell'intervallo è modificata con l'applicazione di una coppia di funzioni di raccordo di durata $\Delta s = b_{ds} - b_{ss}$. La prima funzione ha un valore massimo (in particolare 1) in corrispondenza del bordo sinistro di analisi e valore minimo (in particolare 0) in corrispondenza del punto $b_{sa} + \Delta s$. La seconda funzione ha valore massimo (in particolare 1) in corrispondenza del bordo destro di analisi b_{da} e valore minimo (in particolare 0) in corrispondenza del punto $b_{da} - \Delta s$. Le funzioni di raccordo possono essere del tipo di quelle comunemente utilizzate per questi scopi (p.

es. finestre di Hanning o simili).

A maggior chiarimento dell'invenzione, le figure 4 - 6 riportano alcuni grafici che illustrano l'applicazione del metodo a un segnale fittizio. In tali figure, la parte A mostra tre intervalli consecutivi del segnale originale, con indici $i-1$, i , $i+1$, indicandone il periodo fondamentale p_{ah} ($h = i-1, i, i+1$). Sono indicati i marcatori di periodicità (o di inizio intervallo) M_a e i bordi sinistro e destro di analisi b_{sa} , b_{da} . Le parti B e C mostrano, per ciascun intervallo, rispettivamente la prima e la seconda funzione di raccordo (che nel seguito saranno chiamate per semplicità "funzione A" e "funzione B") e i riferimenti temporali al segnale originale. La parte D mostra la forma d'onda del segnale sintetizzato, come risultato del metodo secondo l'invenzione, con l'indicazione del rispettivo periodo fondamentale p_{sk} ($k = j-1, j, j+1$), dei marcatori di periodicità M_s e dei bordi sinistro e destro di sintesi b_{ss} , b_{ds} . Nella parte E è rappresentata la zona in cui, dopo la traslazione temporale, viene applicato il processo di sovrapposizione e somma delle forme d'onda ottenute con l'applicazione delle due funzioni di raccordo alla parte variabile del segnale originale. Si noti che il numero d'ordine degli intervalli in analisi e in sintesi può essere diverso, in quanto in precedenza possono essersi verificate soppressioni o aggiunte di intervalli.

In particolare, la figura 4 illustra il caso di allungamento del periodo fondamentale (e quindi di diminuzione della frequenza) in sintesi rispetto al segnale originale, in un tratto di segnale in cui non si sono verificate soppressioni o duplicazioni di intervalli. La pesatura è effettuata su ogni intervallo con una rispettiva coppia di funzioni di raccordo. Come conseguenza dell'aumento del periodo, la durata Δs delle funzioni è superiore alla lunghezza della parte variabile del segnale originale, cosicché la funzione B interessa anche l'inizio della forma d'onda relativa all'intervallo successivo, mentre la funzione C interessa una parte di forma d'onda a sinistra del

bordo sinistro di analisi.

Nella fig. 5 è riportata un'analoga rappresentazione nel caso di diminuzione del periodo fondamentale (e quindi aumento della frequenza) in sintesi rispetto al segnale originale. Anche in questo esempio non si sono verificate soppressioni o duplicazioni di intervalli. In questo caso le funzioni B, C interessano un tratto di forma d'onda di durata inferiore al tratto compreso tra b_{sa} e b_{da} .

Infine, la fig. 6 mostra un esempio di allungamento del periodo fondamentale in sintesi nel caso di soppressione di un intervallo del segnale originale (quello con indice i nell'esempio). In sintesi si ottengono due intervalli, indicati con gli indici $j-1$ e j , ottenuti mantenendo, come parte invariante, quella degli intervalli d'indice $i-1$ e $i+1$ nel segnale originale. L'intervallo d'indice $i+1$ nel segnale originale è trattato allo stesso modo dei tre intervalli di segnale originale in fig. 4. La parte modificata dell'intervallo d'indice $j-1$ nel segnale sintetizzato è invece ottenuta sovrapponendo e sommando le due forme d'onda ottenute pesando con la sola funzione B la parte variabile dell'intervallo d'indice $i-1$ nel segnale originale, e pesando con la sola funzione C la parte finale dell'intervallo d'indice i nel segnale originale. In altre parole, la funzione B si applica a destra di b_{sa} nell'intervallo corrente che va riprodotto in sintesi, e la funzione C si applica a sinistra del successivo intervallo da riprodurre. Queste modalità di applicazione delle funzioni di raccordo sono del tutto generali e si applicano anche all'eventuale duplicazione di intervalli e al cambio difono.

A puro titolo esemplificativo, per gli schemi delle figure 4 - 6 si sono utilizzate le seguenti funzioni:

$$0,5 - 0,5 \cdot \cos\{\pi[(\Delta s - 1 + b_{ss} - x_i)/(\Delta s - 1)]^n\} \quad (\text{funzione B})$$

$$0,5 - 0,5 \cdot \cos\{\pi[(x_i - b_{ss})/(\Delta s - 1)]^n\} \quad (\text{funzione C})$$

In queste funzioni, b_{ss} , Δs hanno il significato già visto e sono espresse come

numero di campioni; x_i è il generico campione della parte variabile della forma d'onda originale (con $b_{sa} \leq x_i < b_{sa} + \Delta s$, per la funzione B, e $b_{da} - \Delta s \leq x_i < b_{da}$ per la funzione C); n è un numero che può variare (p. es. da 1 a 3) in dipendenza del rapporto $\Delta s/p_a$: in particolare nel disegno si è considerato $n=1$. Ovviamente, nelle formule, il valore 0,5 può essere sostituito da un generico valore $A/2$ se si utilizza una funzione il cui massimo è A invece di 1, o da una coppia di valori la cui somma è 1 (o A).

Le figure 7A, 7B - 10A, 10B rappresentano alcuni esempi reali di applicazione del procedimento, per due zone del difono "à_m" di fig. 3, utilizzato in due posizioni differenti nella frase dove le regole di sintesi richiedono rispettivamente una diminuzione e un aumento del periodo fondamentale (e quindi un aumento e rispettivamente una diminuzione della frequenza fondamentale). Per tutti gli intervalli sono indicati i marcatori di periodicità, il bordo sinistro di analisi e di sintesi e la frequenza fondamentale, sia in analisi che in sintesi. Le figure con la lettera A riportano la forma d'onda originale e le figure con lettera B il segnale sintetizzato. Le figure 7A, 7B, 8A, 8B mostrano i primi due intervalli del difono in esame (fonema "à") in caso di aumento (figure 7A, 7B) e rispettivamente di diminuzione (figure 8A, 8B) della frequenza fondamentale. Le figure 9A, 9B, 10A, 10B mostrano invece i primi due intervalli del fonema "m" nelle stesse condizioni illustrate nelle figure 7, 8. Per effetto della diminuzione della frequenza, nelle figure 8B e 10B solo il primo intervallo è visibile completamente.

Si descriverà ora una preferita forma di realizzazione del metodo adottato per individuare, per ciascun intervallo da riprodurre in sintesi, il bordo sinistro di analisi e di sintesi. Nell'esempio descritto, si procede in modo differente a seconda che il periodo fondamentale in sintesi sia minore od uguale del periodo di analisi oppure sia maggiore.

La figura 11 è il diagramma di flusso generale delle operazioni compiute nel caso in cui $p_s \leq p_a$.

La prima operazione è il calcolo della funzione ZCR (Zero Crossing Rate) indicativa del numero di passaggi per lo zero (passo 11). In questo calcolo vengono esclusi i passaggi per lo zero che distano dal precedente per meno di un numero limitato di campioni di segnale (ad es. 10), per eliminare oscillazioni non significative del segnale.

Come si vede in fig. 13, ai passaggi per lo zero considerati viene assegnato un indice che va da 1 al descrittore del numero totale LZV (passo 110). Inoltre vengono assegnate le seguenti variabili (passo 111):

- b_{da} (bordo destro di analisi) al valore del periodo di analisi p_a ;
- b_{ds} (bordo destro di sintesi) al valore del periodo di sintesi $b_{da} + \Delta p$;
- $Diff_a_s$ al valore assoluto $|\Delta p|$ della differenza tra il periodo di analisi e quello di sintesi.

In queste relazioni, come pure in quelle esaminate in seguito, i valori del periodo e le lunghezze di certi intervalli sono espressi come numero di campioni.

Tornando alla fig. 11, dopo il calcolo della funzione ZCR si verifica (passo 12) che il numero dei passaggi per lo zero trovato al passo 11 non sia inferiore a una soglia minima di passaggi per lo zero $IndZ_Min$ (p. es. 5 passaggi). Infatti secondo l'invenzione si vogliono riprodurre inalterate, nel segnale sintetizzato, le oscillazioni che seguono immediatamente l'impulso di eccitazione e che, come detto, sono quelle più significative. Se la verifica dà esito positivo, si ricerca un possibile candidato fra i passaggi per lo zero trovati (passo 13) e successivamente si passa a una prima fase di ricerca del bordo sinistro di sintesi e di analisi b_{ss} , b_{sa} (passo 14). Se al termine del passo 14 non si è trovato nessun passaggio per lo zero adatto, si passa a una fase di

continuazione della ricerca (passo 15) e, se anche questa fase non permette ancora di individuare il bordo sinistro di sintesi e di analisi, si passa a una fase di continuazione e conclusione della ricerca (passo 17). Se il confronto al passo 12 indica che il numero di passaggi per lo zero è inferiore alla soglia, si considera arbitrariamente come candidato il passaggio per lo zero di indice $J = \text{IndZ_Min}$ (passo 18) e si esegue una fase di ricerca di b_{sa} e b_{ss} (passo 19) identica a quella compiuta al passo 14: se questa non dà esito, si passa direttamente alla fase 17 di continuazione e conclusione della ricerca, senza passare per la fase 15, per ragioni che risulteranno chiare dalla descrizione di quest'ultima.

Una fase analoga alla fase 17 è prevista anche nel caso di allungamento del periodo fondamentale in sintesi, come si vedrà in seguito. Per semplicità, si è utilizzato uno stesso diagramma di flusso per i due casi, distinguendoli mediante alcune condizioni di ingresso alla fase stessa. In particolare per il caso in esame si pongono le condizioni $r_P \leq 1$ (dove r_P è il rapporto p_s/p_a), $\text{Start} = 0$, $\text{End} = \text{LZV}$, $\text{Step} = +1$ (passo 16 di fig. 11). La prima condizione è evidente. Le altre tre indicano che il ciclo di esame dei passaggi per lo zero previsto nella fase 17 verrà svolto nel senso degli indici crescenti.

Le operazioni compiute nei passi 13-15 e 17 saranno descritte in dettaglio in seguito, con riferimento alle figure 14 - 17.

La figura 12 è il diagramma di flusso generale delle operazioni compiute nel caso in cui il periodo di sintesi p_s è maggiore del periodo di analisi p_a . La prima operazione (passo 21) consiste ancora nel calcolo della funzione ZCR ed è identica a quella indicata con 11 in fig. 11. Successivamente (passo 22) si effettua una ricerca del bordo sinistro di sintesi e di analisi, con modalità che saranno descritte con riferimento alla fig. 18, e, se tale fase non ha esito positivo, si passa a una fase di continuazione e

conclusione della ricerca (passo 24), corrispondente alla fase 17 di fig. 11. Per le operazioni previste nel passo 24 si assegnano le condizioni $r_P > 1$, Start = LZV-1, End = -1, Step = -1. La prima condizione è evidente. Le altre tre indicano che il ciclo di esame dei passaggi per lo zero previsto nella fase 24 verrà svolto in questo caso nel senso degli indici decrescenti.

La fig. 14 mostra il diagramma di flusso delle operazioni di ricerca di un passaggio per lo zero candidato a fungere da bordo sinistro di analisi e di sintesi (passo 13 di fig. 11). Con J si indica l'indice del candidato. In particolare, come candidato si esamina inizialmente il passaggio per lo zero centrale, di indice $J = (LZV+1)/2$ (passo 130) e se ne confronta l'ascissa ZCR(J) con il bordo destro di sintesi (passo 131). Se questo candidato iniziale si trova già a sinistra del bordo destro di sintesi, si passa direttamente alla fase di ricerca del bordo sinistro di analisi e di sintesi (passo 14, fig. 11). In caso contrario si esaminano con un ciclo a ritroso i passaggi per lo zero a sinistra di quello centrale cercando un candidato la cui ascissa sia a sinistra di b_{ds} (passi 132-134). Quando si trova un passaggio per lo zero che soddisfa a tale condizione, lo si considera come candidato (passo 135) e si passa alla fase di ricerca (passo 14 di fig. 1) dopo aver verificato che l'indice del candidato non sia $(LZV+1)/2$ (passo 136). Infatti, dato che si è compiuto un ciclo di ricerca a ritroso perché il candidato iniziale, di indice $(LZV+1)/2$, si trovava a destra di b_{ds} , l'ottenimento di un candidato con tale indice segnala una condizione anomala: se ciò si verifica, si passa alla fase di ricerca dopo aver posto $J=0$. Si procede allo stesso modo se il ciclo termina senza che si sia trovato un candidato.

La fig. 15 mostra le operazioni compiute per la prima fase di ricerca di b_{ss} , b_{sa} (passo 14 di fig. 11). Per tale ricerca si prendono in esame a ritroso i passaggi per lo zero partendo da quello precedente a LZV e si calcola la distanza Diff_z_a tra il bordo

destro di analisi b_{da} e il passaggio per lo zero corrente $ZCR(i)$ (passi 140, 141). Questa distanza, moltiplicata per r_P (rapporto tra il periodo di sintesi p_s e quello di analisi p_a) viene confrontata con $Diff_a_s$ (passo 142), per controllare che vi sia un intervallo temporale sufficiente per l'applicazione della funzione di raccordo. La pesatura con r_P lega la durata di tale funzione all'accorciamento percentuale del periodo ed è volta a garantire un buon raccordo tra gli intervalli. Nel caso in cui $Diff_a_s > Diff_z_a \cdot r_P$, il ciclo di ricerca prosegue (passo 143) fino a quando non si trova un passaggio per lo zero per cui $Diff_a_s \leq (Diff_z_a \cdot r_P)$ o fino all'esaurimento dei passaggi per lo zero: in quest'ultimo caso si esce dal passo 14 e si va al passo 15 (fig. 11) di continuazione della ricerca. Quando la condizione $Diff_a_s \leq Diff_z_a \cdot r_P$ è soddisfatta, l'indice corrente i viene confrontato con l'indice J del candidato (passo 144). Se $i < J$ si continua il ciclo. Se i due indici sono uguali, il passaggio per lo zero corrente viene considerato come bordo sinistro di analisi b_{sa} e di sintesi b_{ss} (passo 147); se invece $i > J$ si calcolano la distanza Δ_a fra il bordo destro di analisi b_{da} e il passaggio per lo zero corrente $ZCR(i)$, la distanza Δ_s fra il bordo destro di sintesi b_{ds} e il passaggio per lo zero corrente $ZCR(i)$ e il rapporto Δ fra Δ_s e Δ_a (passo 145) e si confronta tale rapporto con il valore $(r_P)/2$ (passo 146). Se $\Delta \geq (r_P)/2$ vengono assegnati i bordi sinistri di analisi b_{sa} e di sintesi b_{ss} al passaggio per lo zero corrente (passo 147), altrimenti si passa alla fase 15 (Fig. 11) di continuazione della ricerca. Quest'ultimo confronto indica che non solo si vuole avere una distanza sufficiente tra il bordo sinistro e il bordo destro di sintesi, ma anche che la funzione di raccordo tiene conto dell'accorciamento in sintesi: anche ciò aiuta a ottenere un buon raccordo tra intervalli adiacenti.

La variabile "TRUE" nel passo finale 147 di fig. 14 segnala l'individuazione di b_{sa} e b_{ss} e disabilita le fasi di ricerca successive. La stessa variabile sarà utilizzata

con lo stesso significato anche negli altri diagrammi di flusso relativi alla ricerca del bordo sinistro di analisi e sintesi.

La fase 14 permette di trovare, se esiste, un candidato che stia alla sinistra del bordo destro di sintesi e che sia il più vicino possibile ad esso, garantendo un intervallo temporale sufficiente per l'applicazione della funzione di raccordo: questa fase è il nucleo del criterio di ricerca di b_{sa} e b_{ss} .

Il passo 15 di continuazione della ricerca è illustrato in dettaglio in fig. 16. Tale passo, se viene effettuato (esito negativo della fase 14 e quindi del controllo della condizione TRUE al passo 150), inizia con un nuovo confronto tra LZV e $IndZ_{min}$ (passo 151), destinato questa volta a verificare solo se $LZV > IndZ_{min}$. Se la condizione non è soddisfatta, si passa alla fase 17 di continuazione e conclusione della ricerca. Se $LZV > IndZ_{min}$, si verifica che il passaggio per lo zero di indice $IndZ_{min}$ sia posizionato alla sinistra del bordo destro di sintesi b_{ds} (passo 152) ed in caso positivo si considera tale passaggio come bordo sinistro di analisi b_{sa} e di sintesi b_{ss} (passo 153). Se invece il passaggio per lo zero di indice $IndZ_{min}$ è ancora alla destra del bordo destro di sintesi, si passa alla fase 17 (fig. 11) di continuazione e conclusione della ricerca.

La fase 17 di continuazione e conclusione della ricerca è rappresentata in dettaglio in fig. 17. Dopo il controllo sulla necessità di effettuazione (passo 170), si riprendono in esame i passaggi per lo zero, in ordine di indice crescente. Nel ciclo di esame (passi 171 - 174 di fig. 17) a ogni passo si verifica che il passaggio per lo zero corrente (indicato con Z_{Tmp}) si trovi a sinistra del bordo destro di sintesi b_{ds} e abbia una distanza da questo non inferiore a un valore minimo prefissato δ , ad es. 10 campioni di segnale (passo 173). Se le due condizioni non sono verificate, si esamina il passaggio per lo zero successivo (passo 174), altrimenti si considera

provvisoriamente tale passaggio per lo zero come bordo sinistro di sintesi e di analisi (passo 175) e si continua il ciclo. L'ultimo passaggio per lo zero che soddisfa la condizione 173 sarà considerato come bordo sinistro di sintesi e di analisi (passo 179). Il controllo su r_P previsto al passo 176 è un'ulteriore mezzo di distinzione tra il caso $p_S \leq p_A$ e il caso $p_S > p_A$ e fa sì che i passi 177 e 178 del diagramma di flusso non vengano considerati nel caso in esame.

La fig. 18 illustra la ricerca di b_{SA} e b_{SS} nel caso di allungamento del periodo di sintesi rispetto a quello di analisi. Tale ricerca inizia con un confronto tra l'allungamento in sintesi $Diff_a_s$ e la metà della durata del periodo di analisi p_A (passo 220). Se $Diff_a_s > p_A/2$ si passa direttamente alla fase 24, illustrata in dettaglio in fig. 17. Se $Diff_a_s \leq p_A/2$, si effettua un ciclo di ricerca a ritroso partendo dal passaggio per lo zero precedente a LZV e calcolando la distanza $Diff_z_a$ tra il bordo destro di analisi b_{dA} e il passaggio per lo zero corrente $ZCR(i)$ (passi 221, 222). Questa distanza viene confrontata con $Diff_a_s$ (passo 223): se è minore, il ciclo di ricerca prosegue (passo 224), altrimenti il passaggio per lo zero corrente viene considerato come bordo sinistro di analisi e di sintesi (passo 225). Se al termine del ciclo non si sono potuti determinare b_{SA} e b_{SS} si passa alla fase di continuazione e conclusione della ricerca (passo 24, fig. 12).

Nel caso in cui l'allungamento richiesto in sintesi sia inferiore o uguale a metà del periodo di analisi, le operazioni descritte permettono di trovare, se esiste, un candidato che soddisfa la condizione che la distanza fra il passaggio per lo zero candidato e il bordo destro di analisi sia la prima che supera o eguaglia l'allungamento richiesto.

Nella fase di continuazione e conclusione della ricerca si effettua come detto un ciclo di ricerca a ritroso partendo dal passaggio per lo zero precedente a LZV con le modalità illustrate nei passi 171 - 175 di fig. 17. Inoltre, poiché si considera un

allungamento dell'intervallo (passo 176), per i passaggi per lo zero che soddisfano alle condizioni del passo 173 si calcolano la distanza Δ_a fra il bordo destro di analisi b_{da} e il passaggio per lo zero corrente Z_{Tmp} , la distanza Δ_s fra il bordo destro di sintesi b_{ds} e il passaggio per lo zero corrente Z_{Tmp} e il rapporto Δ tra tali distanze (passo 177). Il rapporto Δ viene confrontato con il doppio del rapporto tra i periodi ($r_P \cdot 2$) per le stesse ragioni viste per il confronto 146 di fig. 15 e, come bordo sinistro di analisi b_{sa} e di sintesi b_{ss} , si prenderà il passaggio per lo zero che soddisfa la condizione $\Delta \leq (r_P \cdot 2)$.

Le condizioni imposte in questa fase permettono di assegnare il bordo sinistro di analisi a un passaggio per lo zero che sta alla sinistra del bordo destro di sintesi, è il più vicino possibile ad esso e inoltre garantisce un intervallo temporale sufficiente per l'applicazione della funzione di raccordo: in particolare, dato un certo periodo in analisi, ad un maggior allungamento richiesto in sintesi corrisponderà un bordo sinistro di analisi posizionato più a ritroso nel periodo originale.

Il metodo descritto può essere messo in opera mediante un convenzionale calcolatore personale o una stazione di lavoro o simili.

E' evidente che quanto descritto è dato unicamente a titolo di esempio non limitativo e che varianti e modifiche sono possibili senza uscire dal campo di protezione dell'invenzione.

CSELT
Centro Studi e Laboratori Telecomunicazioni S.p.A.

Rivendicazioni

1. Procedimento per la sintesi di segnali vocali mediante concatenazione nel tempo di forme d'onda che rappresentano unità elementari di segnale vocale, in cui: si suddividono almeno le forme d'onda associate a suoni vocalizzati in una pluralità di intervalli, corrispondenti alle risposte del condotto vocale a una serie di impulsi di eccitazione delle corde vocali sincroni con la frequenza fondamentale del segnale; si sottopone a pesatura la forma d'onda compresa in ciascun intervallo; si sostituiscono i segnali risultanti dalla pesatura con una loro replica sfalsata nel tempo in misura dipendente da un'informazione prosodica; e si effettua la sintesi mediante sovrapposizione e somma dei segnali sfalsati; caratterizzato dal fatto che:
 - si suddivide un intervallo corrente di segnale originale che deve essere riprodotto in sintesi in una parte invariante, che è compresa tra l'inizio dell'intervallo e un bordo sinistro di analisi, rappresentato da un passaggio per lo zero del segnale vocale originale che soddisfa a condizioni prestabilite, e una parte variabile, che si estende tra il bordo sinistro di analisi e un bordo destro di analisi sostanzialmente coincidente con la fine dell'intervallo corrente, il bordo sinistro e il bordo destro di analisi essendo associati, nel segnale sintetizzato, rispettivamente a un bordo sinistro di sintesi e a un bordo destro di sintesi, dei quali il primo coincide con il bordo sinistro di analisi, con riferimento a un marcatore di inizio intervallo, e il secondo coincide sostanzialmente con la fine dell'intervallo nel segnale sintetizzato;
 - si applica, sulla parte di forma d'onda a destra del bordo sinistro di analisi dell'intervallo corrente di segnale originale, una prima funzione di raccordo, che ha durata uguale a quella del tratto di forma d'onda sintetizzata compreso tra il bordo sinistro e il bordo destro di sintesi e ampiezza progressivamente

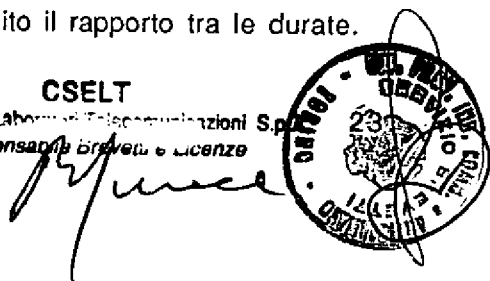
decescente e massima in corrispondenza del bordo sinistro di analisi;

- si applica, sulla parte di forma d'onda a sinistra dell'intervallo immediatamente successivo di segnale originale da riprodurre in sintesi, una seconda funzione di raccordo, che ha durata uguale a quella del tratto di forma d'onda sintetizzata compreso tra il bordo sinistro e il bordo destro di sintesi e ampiezza progressivamente crescente e massima in corrispondenza dell'inizio di detto intervallo successivo;
 - si costruisce ogni intervallo di segnale sintetizzato riproducendo immutata la forma d'onda compresa nella parte invariante dell'intervallo originale, a cui viene unita la forma d'onda ottenuta allineando nel tempo e sommando le due forme d'onda risultanti dall'applicazione delle due funzioni di raccordo.
2. Procedimento secondo la riv. 1, caratterizzato dal fatto che, nel caso in cui la durata di un intervallo venga ridotta o mantenuta invariata per la sintesi rispetto alla durata dell'intervallo corrispondente del segnale originale, il bordo sinistro di analisi e il bordo sinistro di sintesi sono determinati con le operazioni seguenti:
- calcolo del numero di passaggi per lo zero della forma d'onda del segnale originale e assegnazione ai passaggi stessi di un indice crescente dall'inizio verso la fine dell'intervallo;
 - verifica che il numero di passaggi per lo zero sia non inferiore a una prima soglia;
 - ricerca, in caso di esito positivo della verifica, di un passaggio per lo zero candidato a fungere da bordo sinistro di analisi e di sintesi;
 - ricerca a ritroso, tra tutti i passaggi per lo zero dell'intervallo escluso l'ultimo, di un candidato che si trova a sinistra del bordo destro di sintesi, è per quanto possibile vicino ad esso e garantisce un intervallo temporale sufficiente

per l'applicazione delle funzioni di raccordo, e assegnazione del ruolo di bordo sinistro di analisi e di sintesi a tale candidato.

3. Procedimento secondo la riv. 2, caratterizzato dal fatto che in detto calcolo dei passaggi per lo zero non si prendono in considerazione passaggi che hanno una distanza dal passaggio precedente inferiore a una distanza prefissata.
4. Procedimento secondo la riv. 2 o 3, caratterizzato dal fatto che se la ricerca a ritroso ha dato esito negativo e se il numero dei passaggi per lo zero è maggiore della prima soglia, si assegna il bordo sinistro di analisi e il bordo sinistro di sintesi al passaggio per lo zero il cui indice corrisponde a detta soglia, se esso si trova a sinistra del bordo destro di sintesi.
5. Procedimento secondo la riv. 2 o 3, caratterizzato dal fatto che se la ricerca a ritroso ha dato esito negativo e se il numero dei passaggi per lo zero non è maggiore della prima soglia, si effettua un'ulteriore fase di ricerca per individuare i passaggi per lo zero posti a sinistra del bordo destro di sintesi e aventi una distanza da questo non inferiore a una seconda soglia, e si assegna il bordo sinistro di analisi e il bordo sinistro di sintesi al passaggio per lo zero di indice più elevato che soddisfa a dette condizioni.
6. Procedimento secondo la riv. 5, caratterizzato dal fatto che, se il confronto con la prima soglia indica che il numero di passaggi per lo zero è inferiore alla prima soglia, si esegue direttamente detta ricerca a ritroso e, se questa ha esito negativo, si effettua direttamente detta ulteriore fase di ricerca.
7. Procedimento secondo la riv. 1 o 2, caratterizzato dal fatto che, nel caso in cui la durata dell'intervallo venga aumentata per la sintesi rispetto alla durata dell'intervallo corrispondente del segnale originale, il bordo sinistro di analisi e il bordo sinistro di sintesi sono determinati con le operazioni seguenti:

- determinazione del numero di passaggi per lo zero della forma d'onda del segnale originale;
 - confronto tra l'allungamento della durata dell'intervallo per la sintesi e la durata dell'intervallo originale, per verificare che l'allungamento non superi metà della durata dell'intervallo originale;
 - in caso di esito positivo della verifica, ricerca a ritroso, tra tutti i passaggi per lo zero escluso l'ultimo, di un passaggio per lo zero candidato che si trovi alla sinistra del bordo destro di sintesi e sia il primo per cui la distanza dal bordo destro di sintesi non è inferiore all'allungamento della durata dell'intervallo, il bordo sinistro di analisi e il bordo sinistro di sintesi essendo assegnati all'eventuale passaggio per lo zero che soddisfa a detta condizione.
8. Procedimento secondo la riv. 7, caratterizzato dal fatto che in detto calcolo dei passaggi per lo zero non si prendono in considerazione passaggi che hanno una distanza dal passaggio precedente inferiore a una distanza prefissata.
9. Procedimento secondo la riv. 7 o 8, caratterizzato dal fatto che, se l'allungamento della durata dell'intervallo supera metà della durata dell'intervallo originale o se la ricerca a ritroso non ha esito, si effettua un'ulteriore fase di ricerca a ritroso per individuare i passaggi per lo zero posti a sinistra del bordo destro di sintesi e aventi una distanza da questo non inferiore a una terza soglia; per tali passaggi per lo zero si calcolano le distanze dal bordo destro di sintesi e dal bordo destro di analisi e il rapporto tra tali distanze; si confronta tale rapporto con il valore del rapporto tra la durata dell'intervallo in sintesi e la durata dell'intervallo originale e si assegna il bordo sinistro di analisi e il bordo sinistro di sintesi al passaggio per lo zero di indice più basso tra quelli per cui il rapporto tra le distanze dai bordi non superi di un fattore prestabilito il rapporto tra le durate.



4A000075C

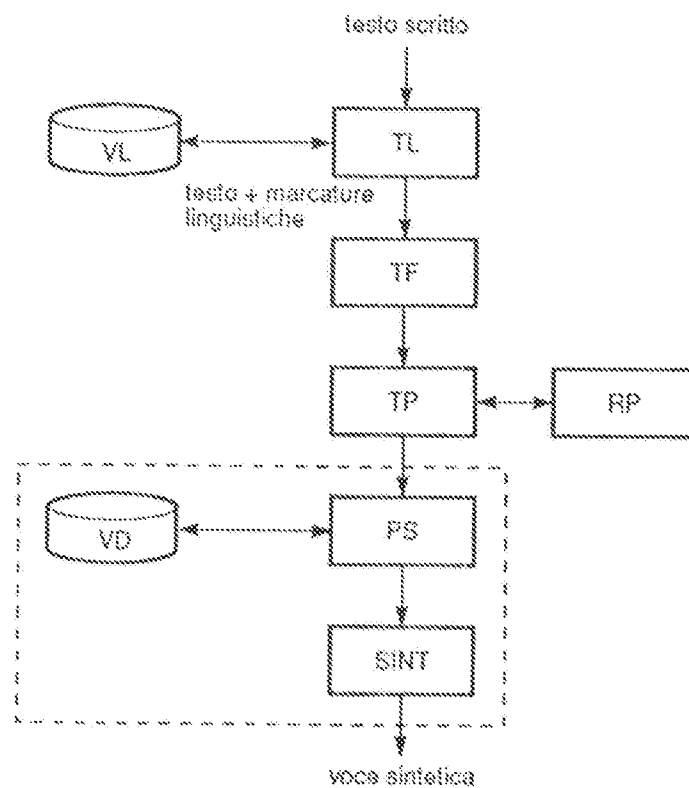


FIG. 1

COELT
Centro Studi e Laboratori Telecomunicazioni S.p.A.
Il Responsabile Brevetti e Licenze

[Handwritten signature]



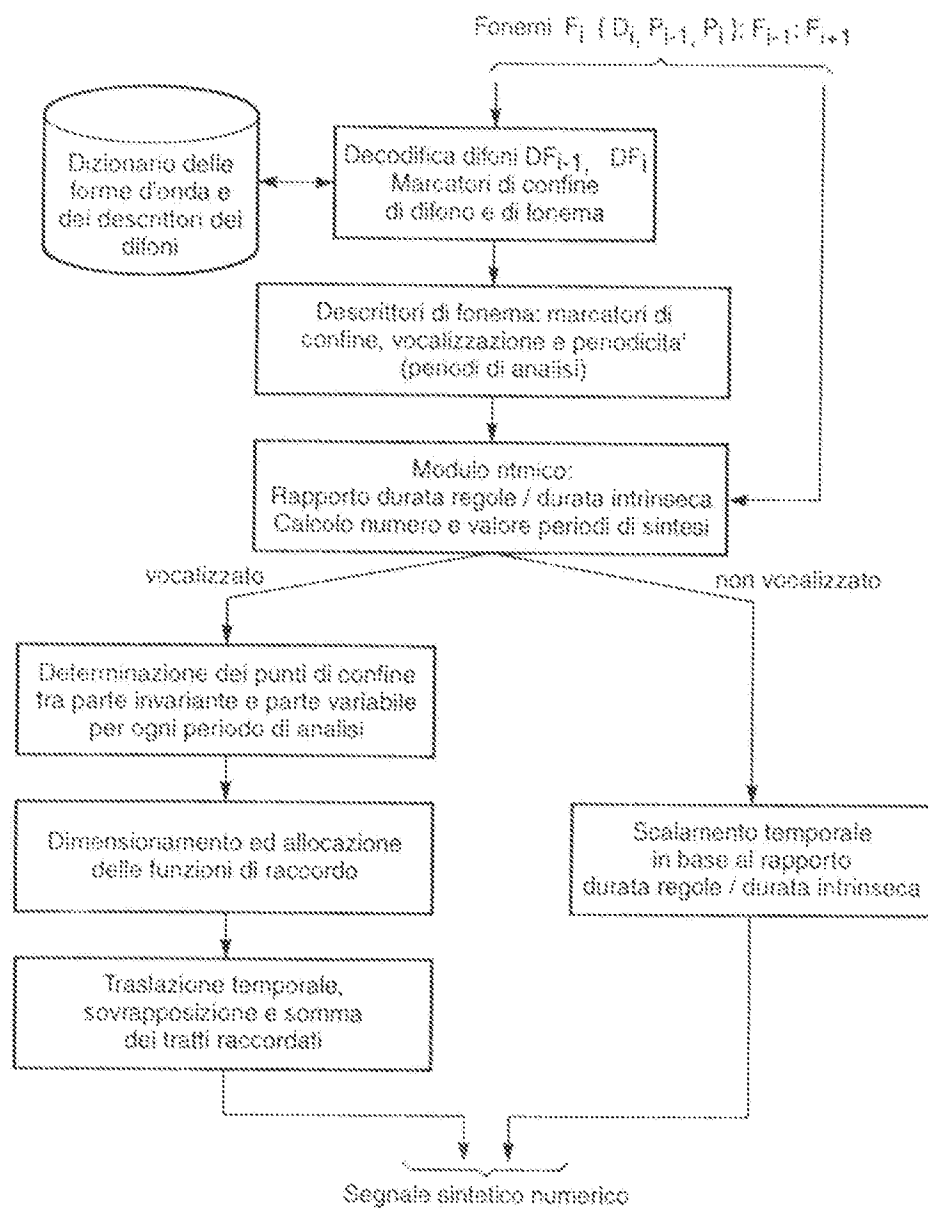
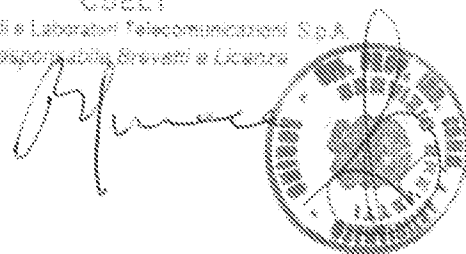


FIG. 2

CSELT
Centro Studi e Laboratori Telecomunicazioni S.p.A.
Il Responsabile Brevetti e Licenze



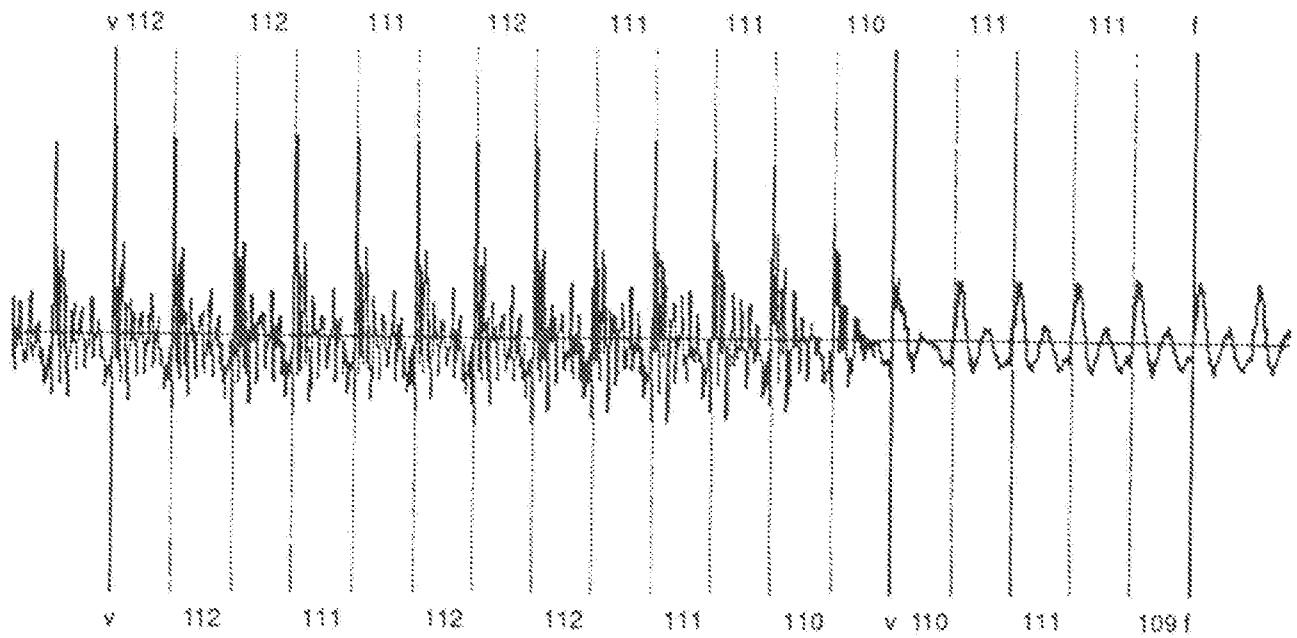
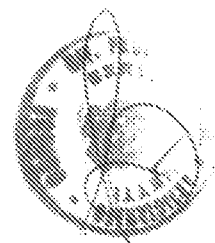


FIG. 3

CSELT
 Centro Studi e Laboratori Telecomunicazioni S.p.A.
 il Responsabile Diritto e Licenze

[Handwritten signature]



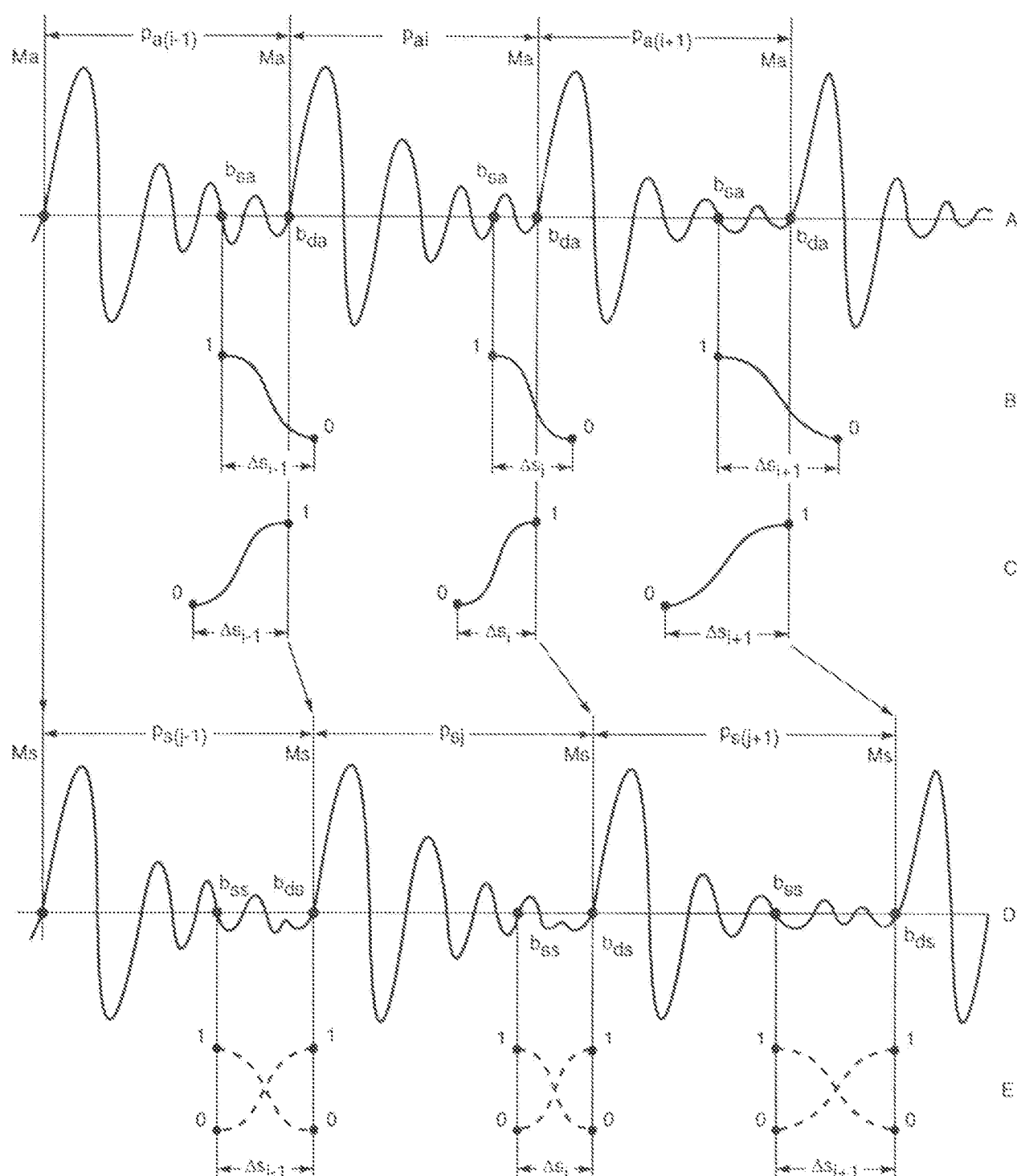


FIG. 4

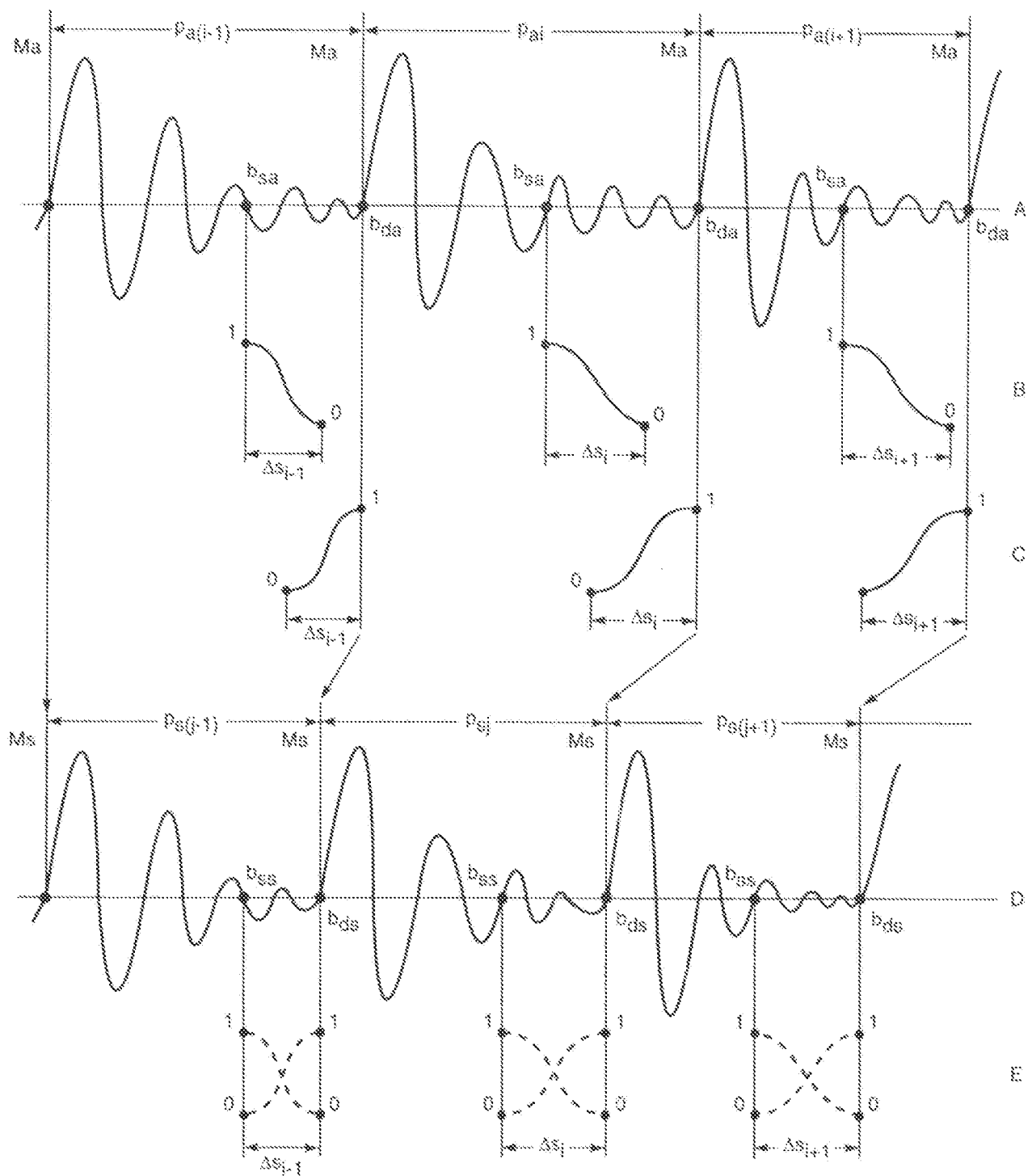
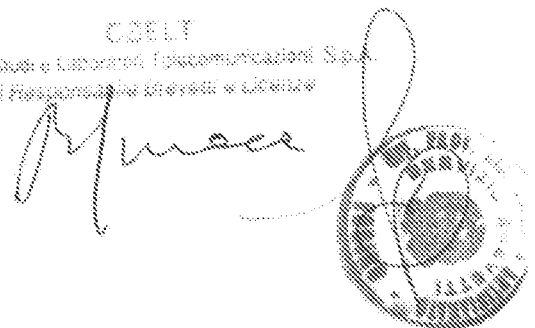


FIG. 5

GSELT
 Centro Studi e Innovazioni Telecomunicazioni S.p.A.
 Il Responsabile Brevetti e Licenze



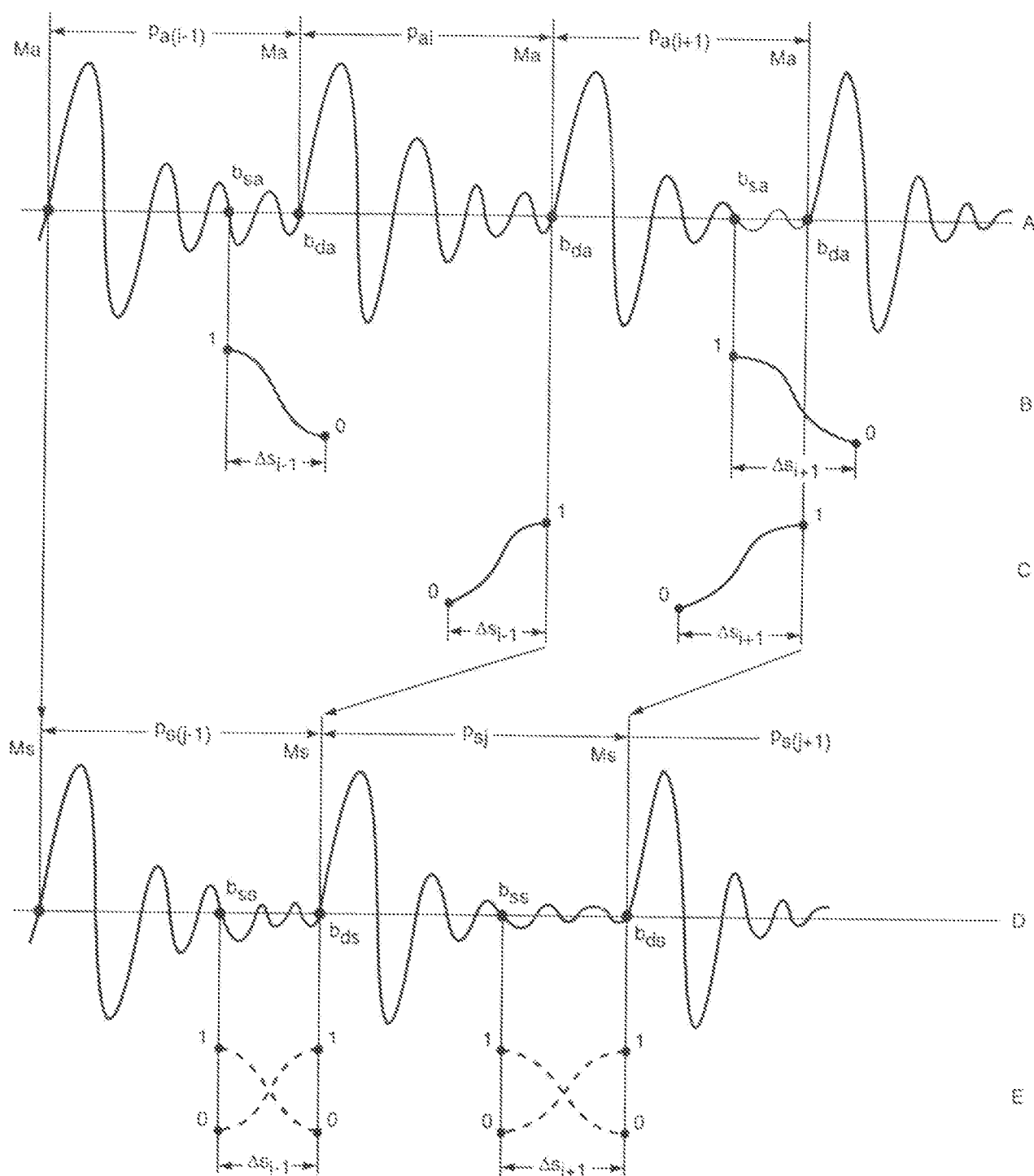


FIG. 6

CSELT
 Centro Studi e Laboratori Telecomunicazioni S.p.A.
 Il Plesso di via Broletto 15 - 10121 Torino



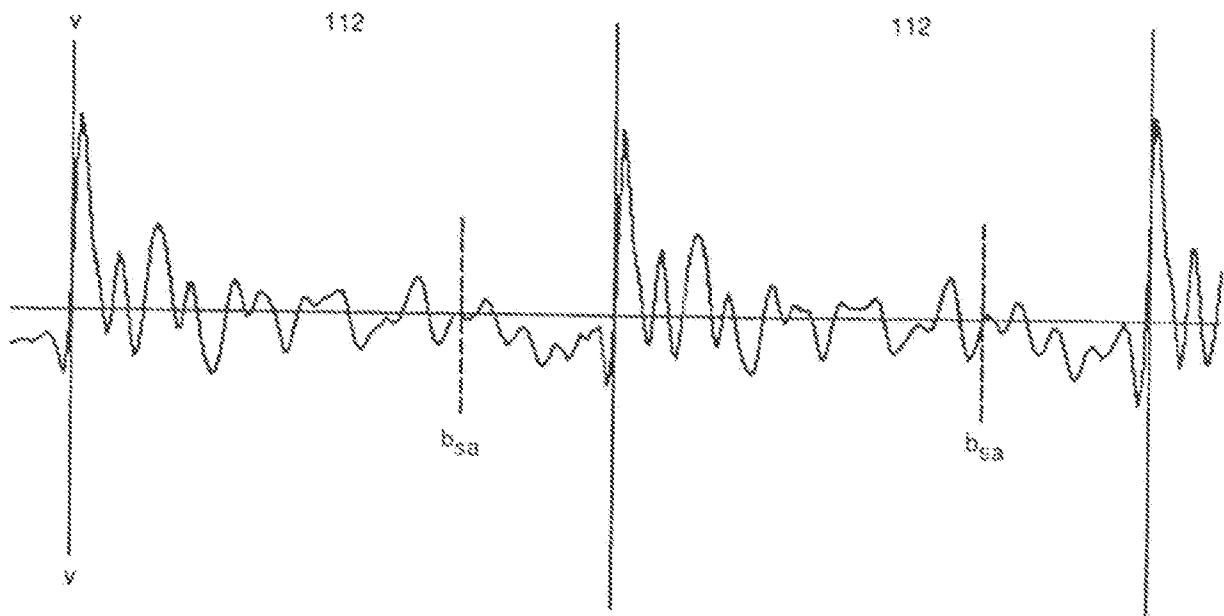


FIG. 7A

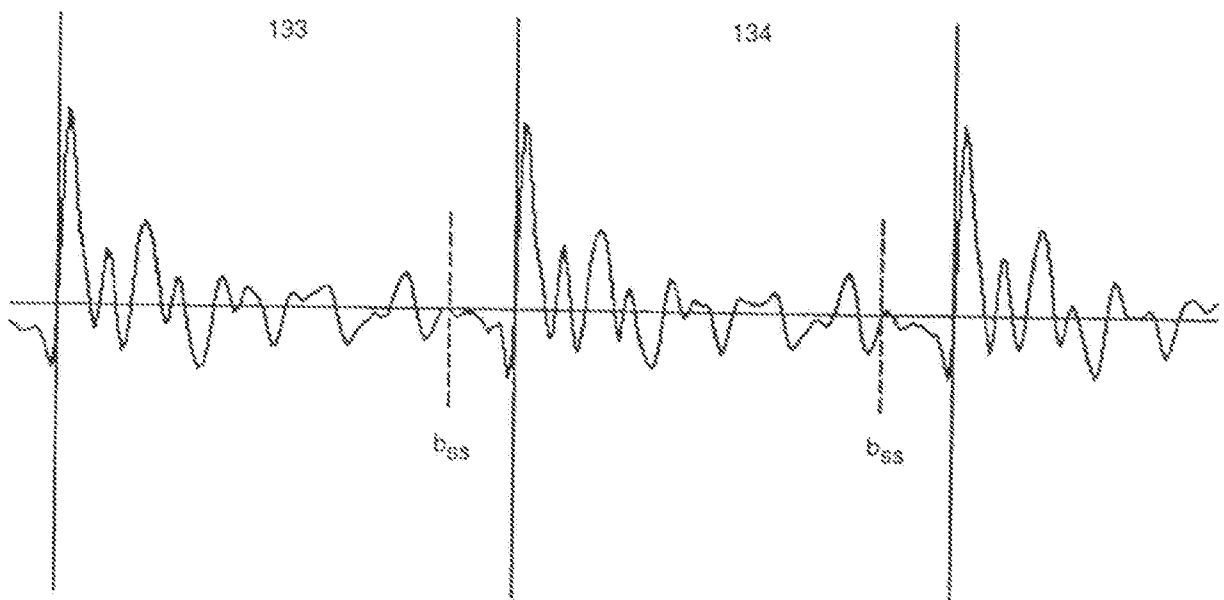


FIG. 7B

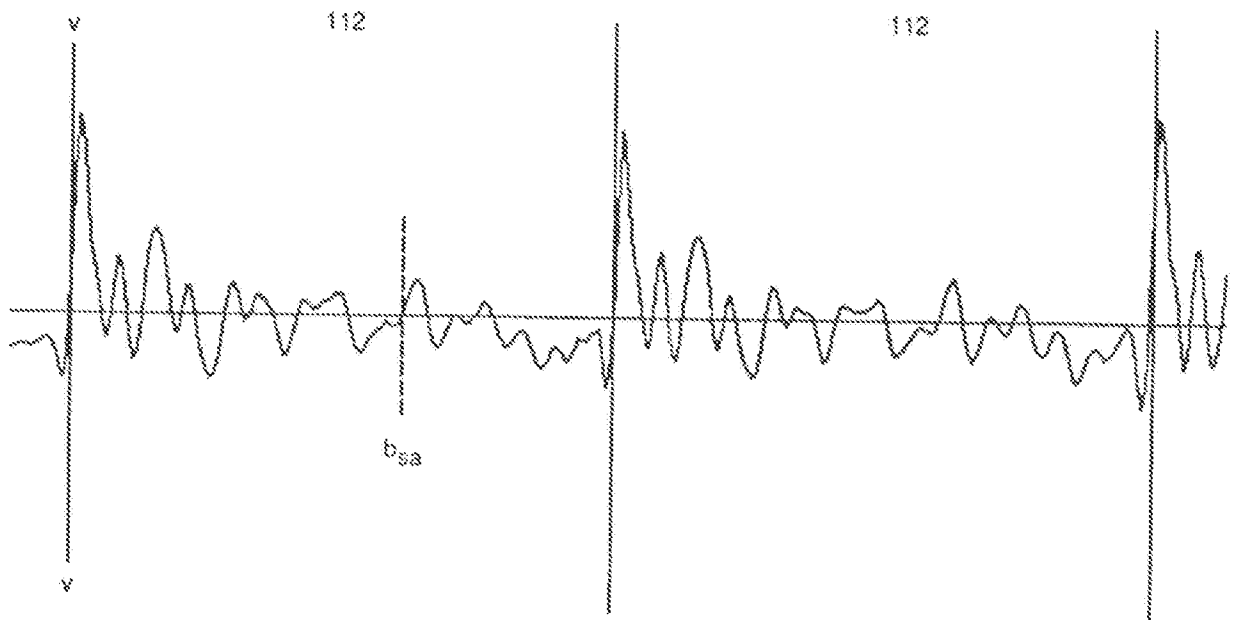


FIG. 8A

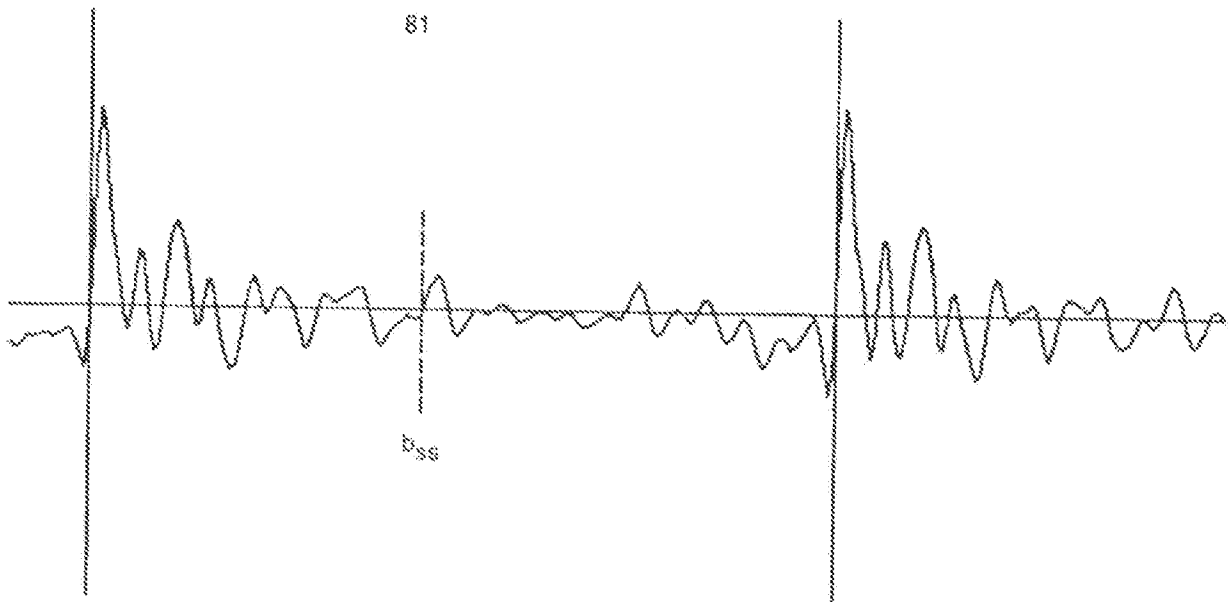
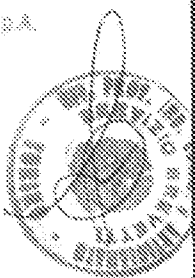
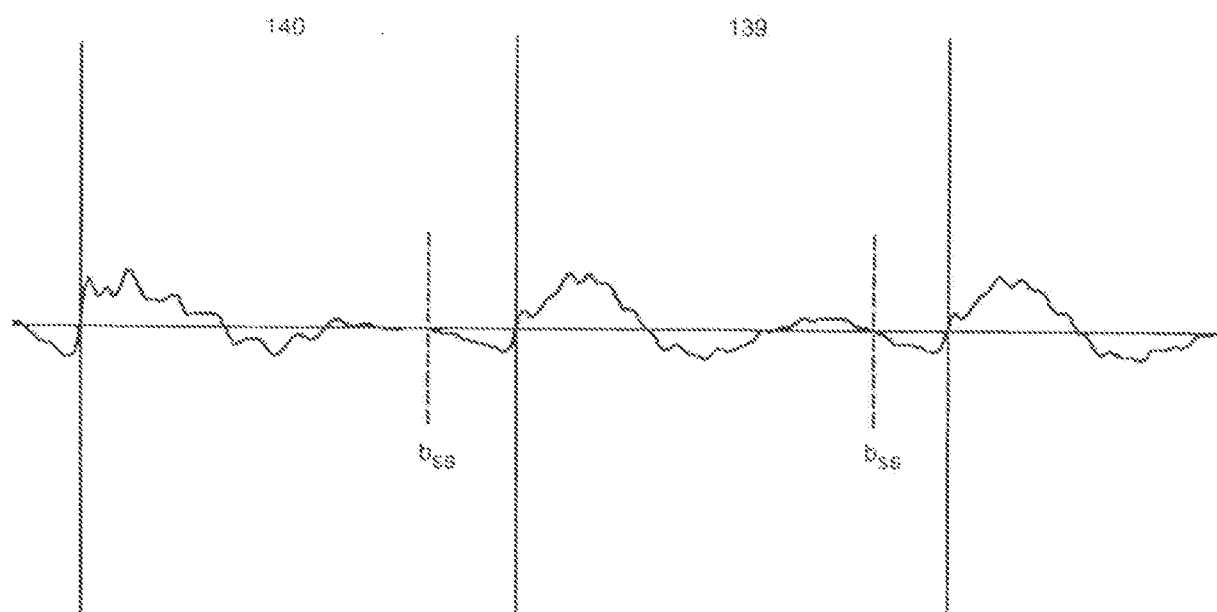
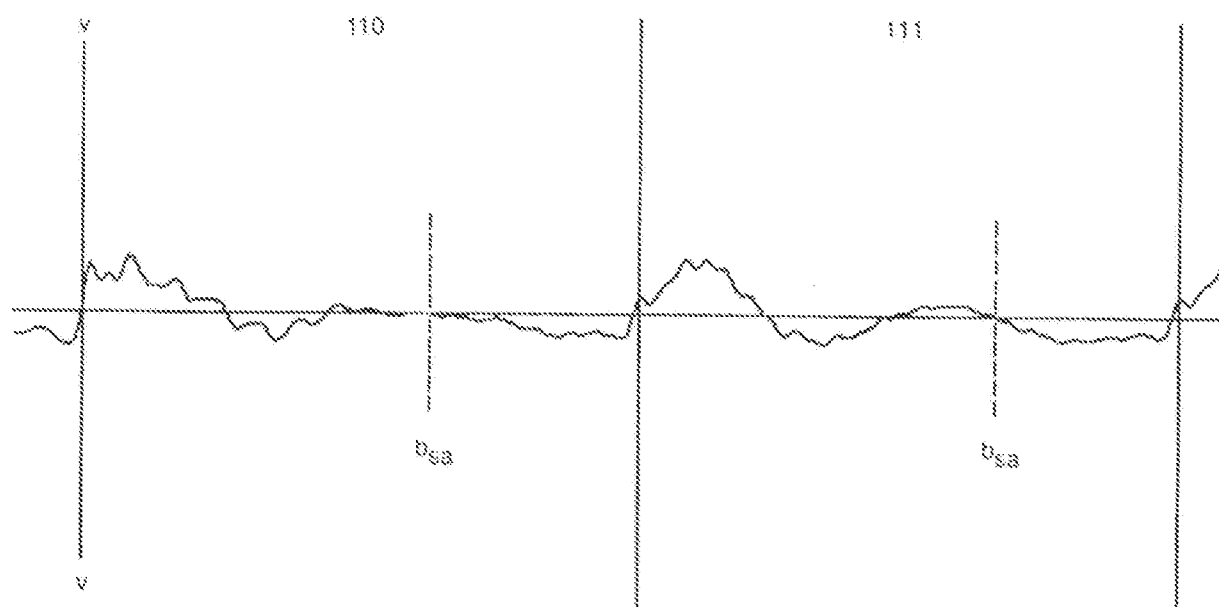


FIG. 8B

CSELT
 Centro Studi e Ricerche Telecomunicazioni S.p.A.
 Il Fascicolo dei Brevetti e Licenze

Foruncce





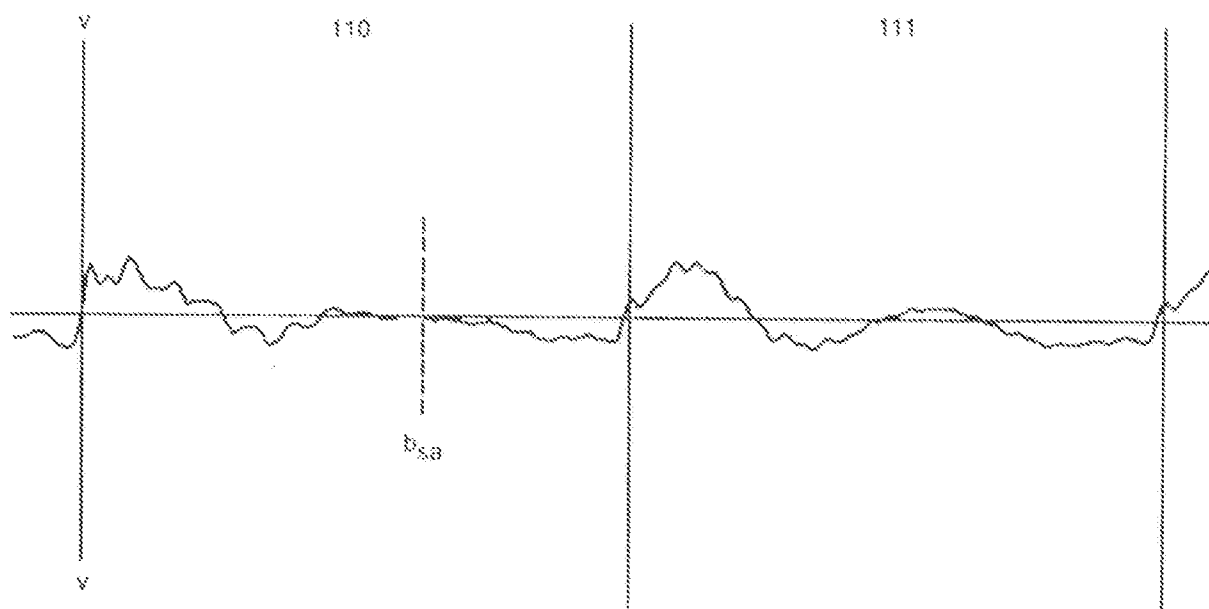


FIG. 10A

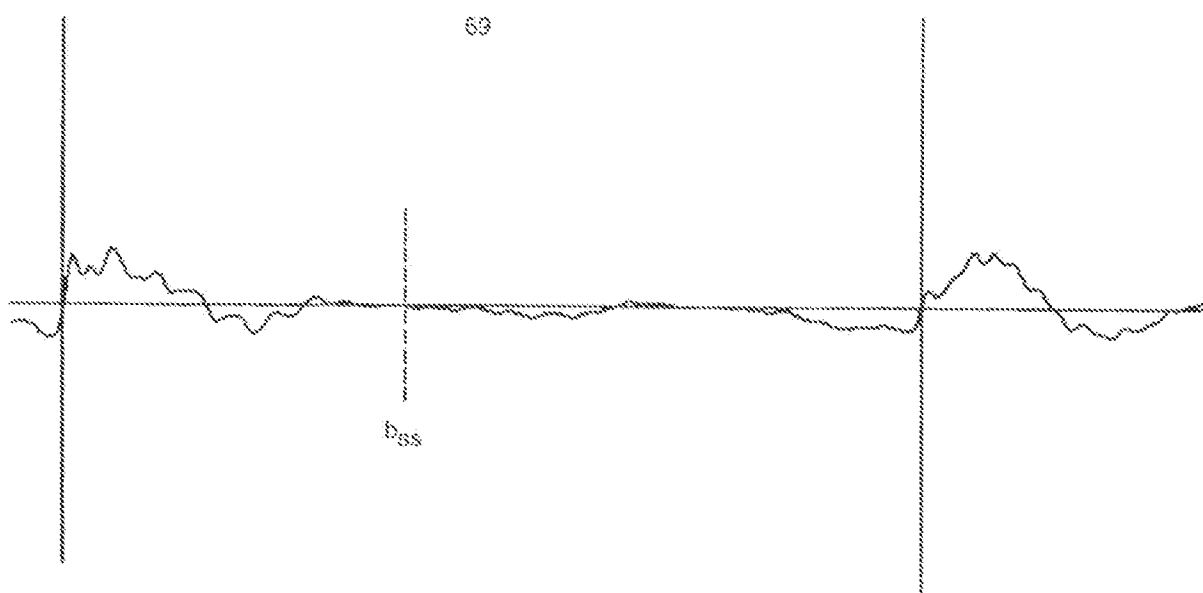
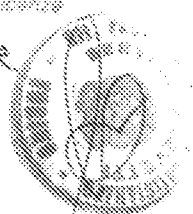


FIG. 10B

[Handwritten signature]



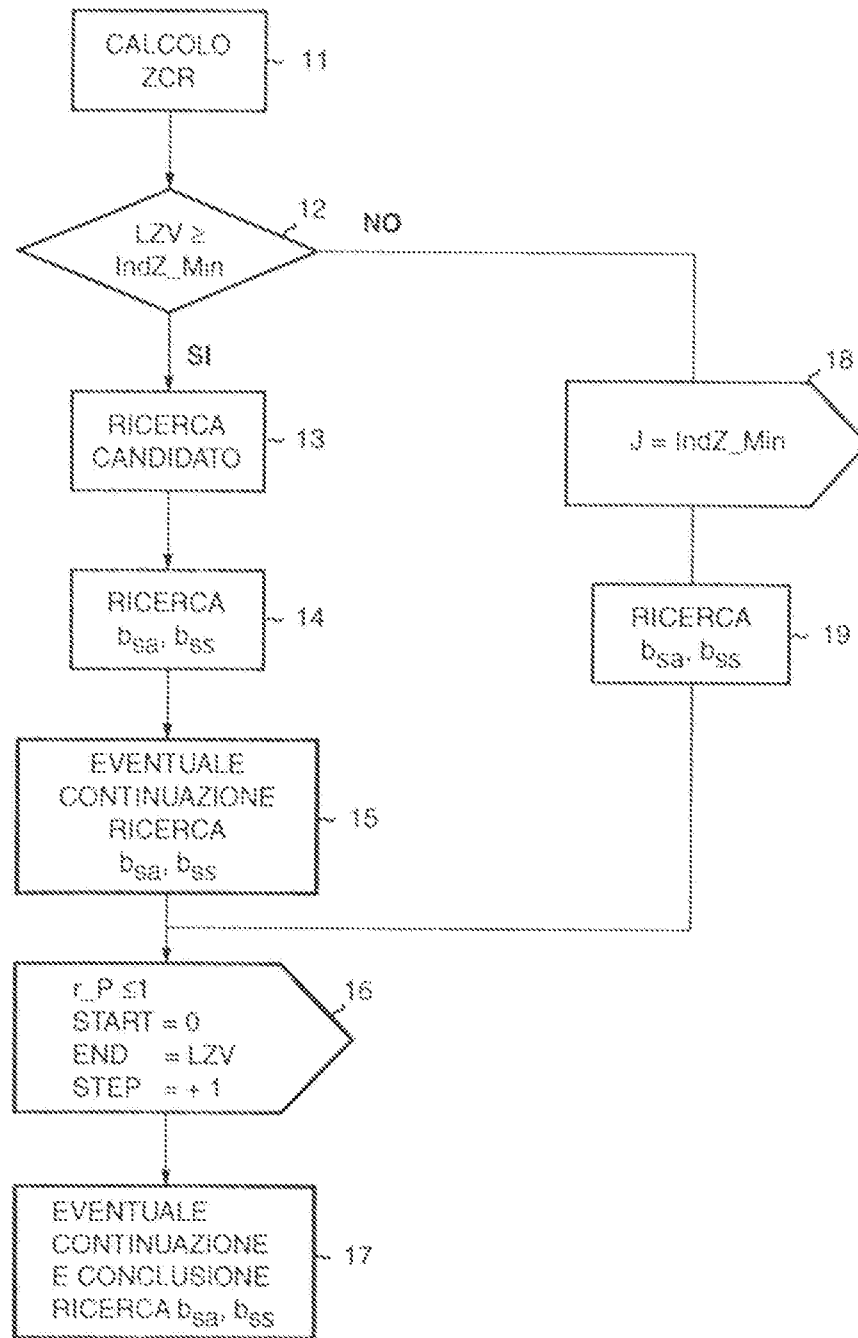
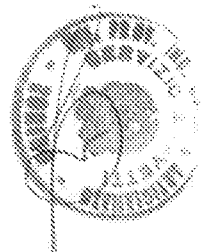


Fig.11



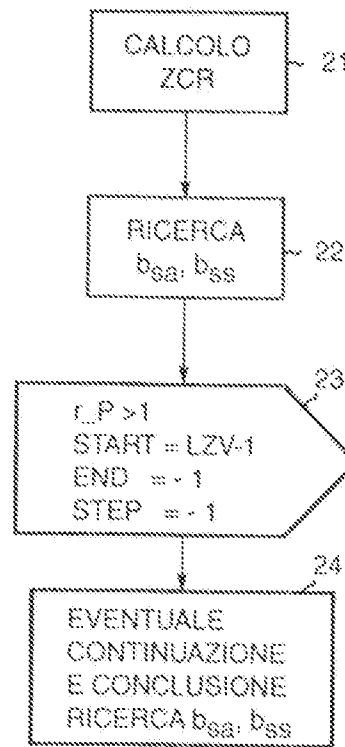


Fig.12

11 (21)

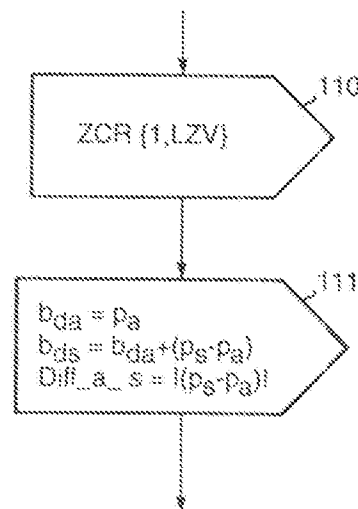
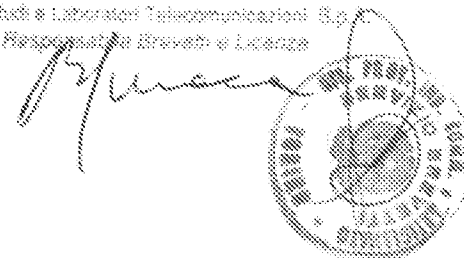


Fig.13

CSELT
Centro Studi e Laboratori Telecomunicazioni S.p.A.
Il Responsabile Brevetti e Licenze



TO 94A00075G

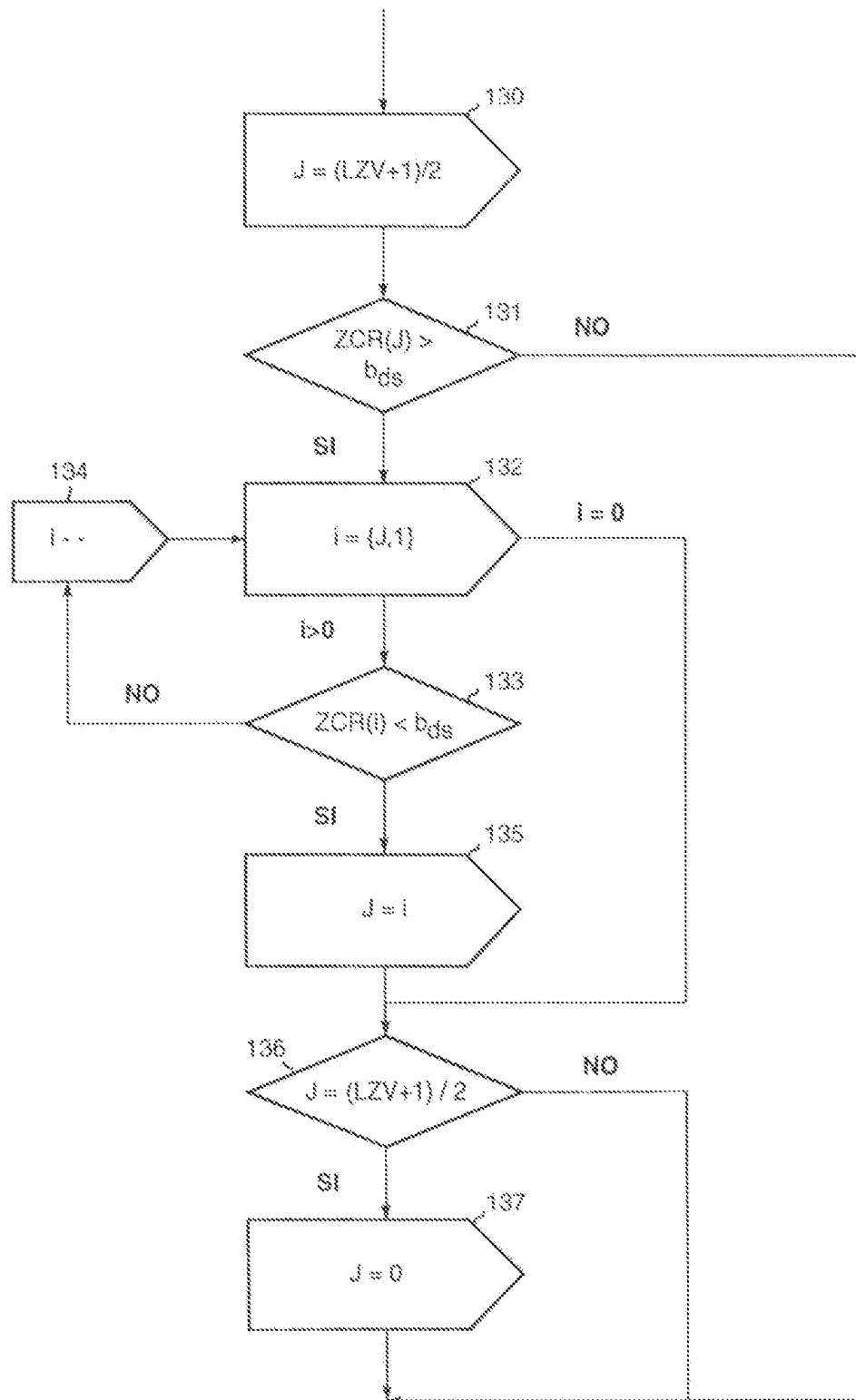
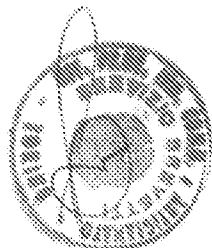


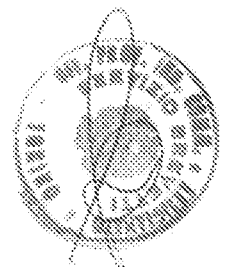
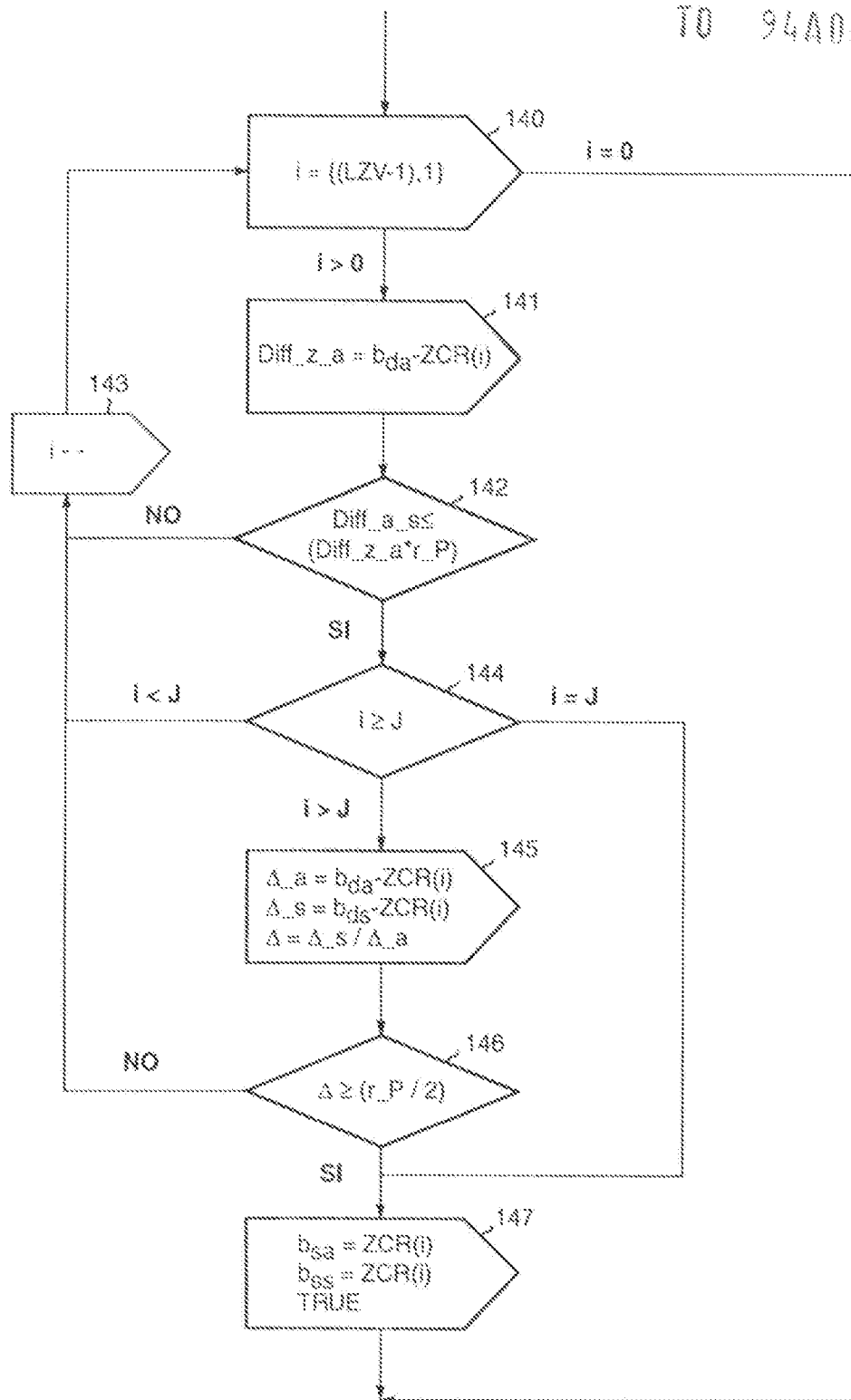
Fig.14



CSELT
Centro Studi e Ricerche Telecomunicazioni S.p.A.
Brevetto e Licenza

Handwritten signature

TO 94A000753



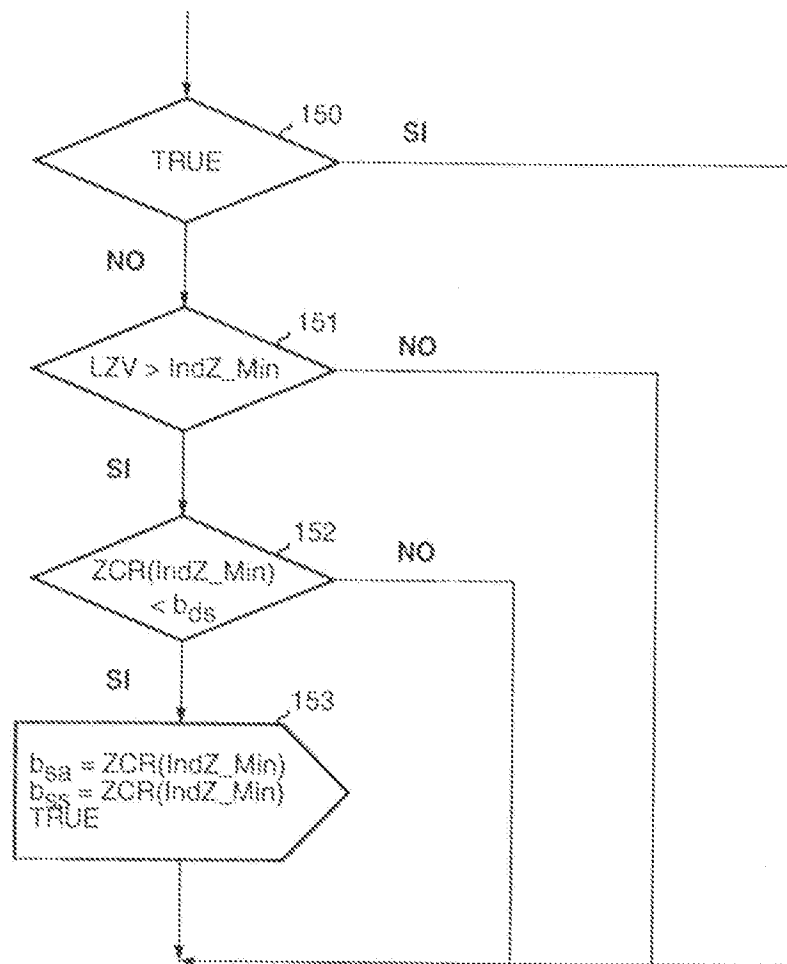


Fig.16



[Handwritten signature]

SI TO 94A000075C

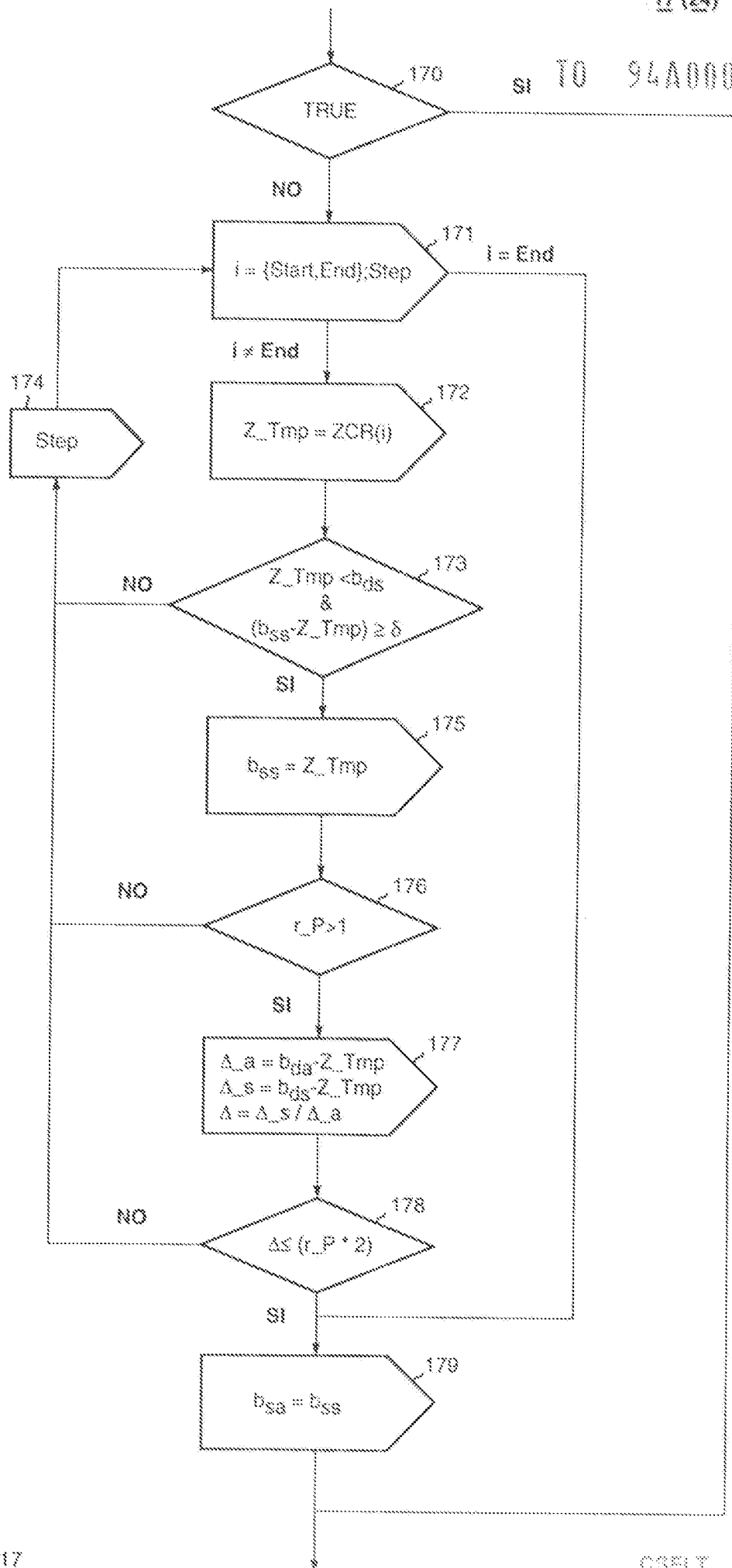
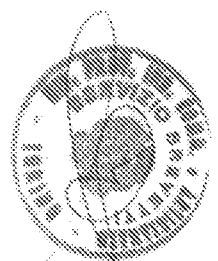


Fig. 17



R. M. M.