



(12) PATENT

(19) NO

(11) 323734

(13) B1

NORGE

(51) Int Cl.

G10L 19/00 (2006.01)

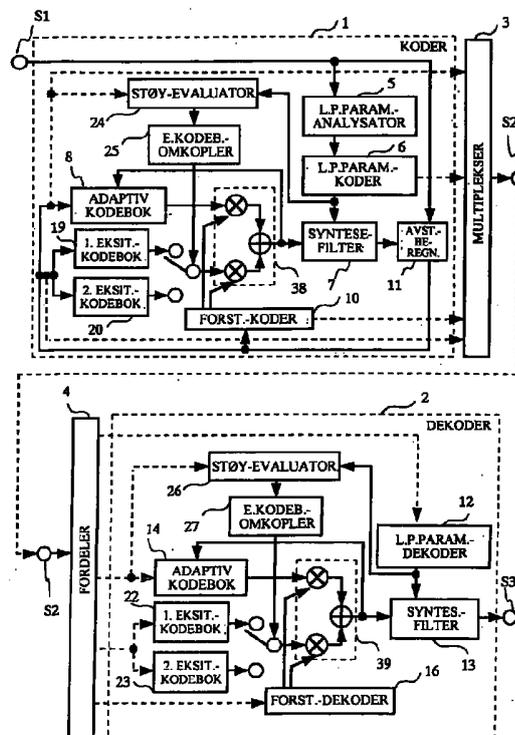
G10L 19/14 (2006.01)

Patentstyret

(21)	Søknadsnr	20035109	(86)	Int.inng.dag og søknadsnr	1998.12.07 PCT/JP98/05513
(22)	Inng.dag	2003.11.17	(85)	Videreføringssdag	2003.11.17
(24)	Løpedag	1998.12.07	(30)	Prioritet	1997.12.24, JP, 9-354754
(41)	Alm.tilgj	2000.06.23			
(45)	Meddelt	2007.07.02			
(62)	Avdelt fra	20003321, med inndato 2000.06.23			
(73)	Innehaver	Mitsubishi Denki KK, 7-3, Marunouchi 2-chome, Chiyoda-ku, 100-8310 TOKYO, JP			
(72)	Oppfinner	Tadashi Yamaura, c/o Mitsubishi Denki KK, 2-3, Marunouchi 2-chome, Tokyo 100-8310, JP			
(74)	Fullmektig	Bryn Aarflot AS, Postboks 449 Sentrum, 0104 OSLO			

(54)	Benevnelse	Fremgangsmåte for talekoding, fremgangsmåte for taledekoding, samt deres apparater
(56)	Anførte publikasjoner	JP 8185198
(57)	Sammendrag	

Høykvalitetstale reproduseres ved hjelp av en forholdsvis liten datamengde ved koding og dekoding av tale for å utføre kompresjonskoding og -dekoding av et talesignal til et digitalsignal. Ved en fremgangsmåte for talekoding i samsvar med kodeksitert lineær prediksjon (CELP) ved talekoding, blir støynivået for en tale innenfor en bestemt kodingsperiode evaluert ved bruk av én kode eller et kodingsresultat for minst én av parametrene spektruminformasjon, effektinformasjon og tonehøydeinformasjon, idet forskjellige eksiteringskodebøker anvendes basert på et evalueringsresultat.



Teknisk område

Denne oppfinnelse gjelder fremgangsmåter for talekoding og dekoding av tale samt apparater for slik talekoding og dekoding ved utførelse av kompresjonskoding og -dekoding av et talesignal til et digitalsignal. Spesielt gjelder oppfinnelsen en fremgangsmåte for talekoding, en fremgangsmåte for taledekoding samt apparater for henholdsvis talekoding og taledekoding med det formål å gjengi tale med høy kvalitet og ved lave bit-takter.

Bakgrunnsteknikk

Innenfor den beslektede teknikk er kodeeksitert lineær prediksjon (Code-Excited Linear Prediction: CELP) velkjent som en effektiv talekodingsmetode, og dens teknikk er beskrevet i "Code-excited linear prediction (CELP): Highquality speech at very low bit rates", ICASSP 1985, sidene 937-940, av M. R. Schroeder og B. S. Atal i 1985.

Et annet eksempel for et system og en fremgangsmåte relatert til talekoding og -dekoding, der flere eksiteringskodebøker kobles inn på basis av en tonefrekvens som velges i en adaptiv kodebok, fremlegges i JP8,185,198.

Fig. 6 viser et eksempel på en fullstendig konfigurasjon av en CELP-talekodings- og taledekodingsmetode. I fig. 6 er det vist en koder 101, en dekoder 102, multipleksingsutstyr 103 og divisjonsutstyr 104.

Koderen 101 omfatter en innretning for lineær prediksjonsparameter-analysering 105, en innretning for lineær prediksjonsparameter-koding 106, et syntesefilter 107, en adaptiv kodebok 108, en eksiteringskodebok 109, en innretning 110 for forsterkningskoding, en innretning for avstandsberegning 111, samt en innretning for avveiling og addering 138. Dekoderen 102 omfatter en innretning for prediksjonsparameter-dekoding 112, et syntesefilter 113, en adaptiv kodebok 114, en eksiteringskodebok 115, en innretning for forsterkningsdekoding 116, og en innretning for avveiling og addering 139.

Ved CELP-talekoding blir tale innenfor en tidsramme på omkring 5-50 ms oppdelt i spektruminformasjon og eksiteringsinformasjon, samt kodet.

Det vil nå bli forklart hvorledes fremgangsmåten for CELP-talekodingen virker. I koderen 101 analyserer innretningen for lineær prediksjonsparameter-analyse en inngangstale S101 og trekker ut en lineær prediksjonsparameter som

utgjør spektruminformasjon for talen. Innretningen 106 for lineær prediksjonsparameter-koding koder denne lineære prediksjonsparameter og fastlegger en kodet lineær prediksjonsparameter som en koeffisient for syntesefilteret 107.

Kodingen av eksiteringsinformasjonen vil nå bli forklart.

5 Et tidligere eksiteringssignal er lagret i den adaptive kodebok 108. Denne adaptive kodebok 108 avgir en tidsserievektor som tilsvarer en adaptiv kode som er tilført fra avstandsberegneren 111, og som er generert ved å gjenta det tidligere eksiteringssignal periodisk.

Flere tidsserievektorer som er utviklet ved reduksjon av taleforvrengning
10 med henblikk på opplæring og den tilsvarende kodede tale er f.eks. lagret i eksiteringskodeboken 109. Denne eksiteringskodebok 109 avgir en tidsserievektor som tilsvarer en eksiteringskode tilført fra avstandsberegneren 111.

Hver av de tidsserievektorer som avgis fra den adaptive kodebok 108 og eksiteringskodeboken 109 avveies ved bruk av en tilsvarende forsterkning som
15 frembringes av innretningen 110 for forsterkningskoding og adderes ved hjelp av innretningen 138 for avveiling og addering. Et addisjonsresultat blir så frembrakt og overføres til syntesefilteret 107 som eksiteringssignaler, og det dannes da en kodet tale. Innretningen for avstandsberegning 111 beregner en avstand eller forskjell mellom den kodede tale og inngangstalen S101 og søker etter en adaptiv
20 kode, eksiteringskode og forsterkninger for å redusere denne forskjell til et minimum. Når den ovenfor omtalte koding er avsluttet, blir en lineær prediksjonsparameter-kode og den adaptive kode, eksiteringskoden og forsterkningskodene for å nedsette en forvrengning mellom inngangstalen og den kodede tale avgitt som et kodingsresultat.

25 Det vil nå bli forklart hvorledes fremgangsmåten for CELP-taledekodingen virker.

I dekoderen 102 dekoder innretningen 112 for lineær prediksjonsparameter-dekoding den lineære prediksjonsparameterkode for den lineære prediksjonsparameter, samt fastlegger den lineære prediksjonsparameter
30 som en koeffisient for syntesefilteret 113. Den adaptive kodebok 114 avgir en tidsserievektor som tilsvarer en adaptiv kode, og som genereres ved å gjenta et tidligere eksiteringssignal periodisk. Eksiteringskodeboken 115 avgir en tidsserievektor som tilsvarer en eksiteringskode. Disse tidsserievektorer avveies ved bruk av til-

svarende forsterkninger, som da dekodes fra forsterkningskodene ved hjelp av innretningen 116 for forsterkningsdekoding, samt adderes ved hjelp av innretningen 139 for avveiling og addering. Det frembringes da et addisjonsresultat som avgis til syntesefilteret 113 som et eksitasjonssignal, og en utgangstale S103 frembringes på denne måte.

Blant de foreliggende metoder for CELP-talekoding og taledekoding er en forbedret fremgangsmåte for talekoding og -dekoding med henblikk på å reproducere høykvalitetstale i samsvar med beslektet teknikk beskrevet i "Phonetically - based vector excitation coding of speech at 3.6 kbps," ICASSP 1989, sidene 49-52, av S. Wang og A. Gersho i 1989.

Fig. 7 viser et eksempel på en fullstendig konfigurasjon av fremgangsmåten for talekoding og taledekoding i henhold til den beslektede teknikk, og samme henvisningstegn er her anvendt som for de viste innretninger i fig. 6.

I fig. 7 omfatter koderen 101 en bestemmelsesinnretning 117 for taletilstanden, en omkoplingsinnretning 118 for eksiteringskodebok, en første eksiteringskodebok 119 og en andre eksiteringskodebok 120. Dekoderen 102 omfatter en omkoplingsinnretning 121 for eksiteringskodebok, en første eksiteringskodebok 122 og en andre eksiteringskodebok 123.

Det vil nå bli forklart hvorledes fremgangsmåten for koding og dekoding virker i denne konfigurasjon. I koderen 101 analyserer bestemmelsesinnretningen 117 for taletilstanden inngangstalen S101, og bestemmer om talen befinner seg i én av to tilstander, f.eks. stemt eller ustemt. Omkoplingsinnretningen 118 for eksiteringskodebok kopler inn den av eksiteringskodebøkene som skal anvendes ved kodingen på grunnlag av resultatet av en taletilstandsbestemmelse. Hvis f.eks. talen er stemt, anvendes da den første eksiteringskodebok 119, og hvis talen er ustemt, så vil den andre eksiteringskodebok 120 bli brukt. Omkoplingsinnretningen 118 for eksiteringskodebok koder så hvilken eksiteringskodebok som er anvendt ved kodingen.

I dekkeren 102 kopler omkoplingsinnretningen 121 for eksiteringskodebok om mellom den første eksiteringskodebok 122 og den andre eksiteringskodebok 123 basert på en kodeanvisning som angir hvilken eksiteringskodebok som er blitt anvendt i koderen 101, slik at den eksiteringskodebok som er blitt anvendt i koderen 101 også anvendes i dekkeren 102. I samsvar med denne konfigurasjon ut-

nyttes således eksitasjonskodebøker som til enhver tid er egnet for koding i de forskjellige taletilstander, og disse eksiteringskodebøker koples inn på grunnlag av en tilsvarende tilstand av en inngangstale. På denne måte kan tale gjengis med høy kvalitet.

5 En fremgangsmåte for talekodingen og -dekoding som går ut på omkopling mellom flere eksiteringskodebøker uten økning av overføringsbitantallet og i samsvar med den beslektede teknikk er angitt i japansk ugransket offentliggjort patent-søknad 8-185198. Disse flere eksiteringskodebøker koples inn på grunnlag av en tonefrekvens som velges i en adaptiv kodebok, og en eksiteringskodebok som er
10 egnet for å angi egenskaper ved en inngangstale kan da anvendes uten at overføringsdata økes.

Som angitt for den fremgangsmåte for talekoding og taledekoding som er vist i fig. 6 i samsvar med kjent beslektet teknikk, anvendes en enkelt eksiteringskodebok for å frembringe en syntetisk tale. Støyfrie tidsserievektorer med mange
15 pulser bør lagres i eksiteringskodeboken for å frembringe en kodet tale av høy kvalitet selv ved lave bit-takter. Når det da foreligger tale med støy, f.eks. bakgrunnsstøy, frikative konsonanter, etc., kodes og syntetiseres, vil det være et problem at en kodet tale frembringer en unaturlig lyd, f.eks. "Jiri-Jiri" og "Chiri-Chiri". Dette problem kan løses, hvis eksiteringskodeboken omfatter bare støy-
20 tidsserievektorer. I dette tilfelle vil imidlertid kvaliteten av den kodede tale bli degradert i sin helhet.

Ved den forbedrede fremgangsmåte for talekoding og taledekoding som er vist i fig. 7 og er i samsvar med beslektet tidligere teknikk, blir da de flere eksiteringskodebøker koplet inn på grunnlag av inngangstalens tilstand for derved å
25 frembringe en kodet tale. Det vil da være mulig å bruke en eksiteringskodebok som omfatter støy-tidsserievektorer innenfor en ustemt støyperiode av inngangstalen samt en eksiteringskodebok som omfatter støyfrie tidsserievektorer innenfor en stemt periode som er forskjellig fra den utstemte støyperiode, er angitt som et eksempel. Selv om en tale med støy blir kodet og syntetisert vil da ikke noen u-
30 naturlig lyd, f.eks. "Jiri-Jiri", blir frembrakt. Da imidlertid den eksiteringskodebok som anvendes ved kodingen også brukes ved dekodingen, vil det være nødvendig å kode og overføre data som angir hvilken eksiteringskodebok som er brukt. Dette blir da en hindring for å oppnå lave bit-takter.

I henhold til fremgangsmåten for talekoding og -dekoding ved omkopling mellom flere eksiteringskodebøker uten å øke en overførings bitantall i samsvar med den beslektede teknikk, blir eksiteringskodebøkene koplet om på grunnlag av en tonehøydeperiode valgt i den adaptive kodebok. Den tonehøydeperiode som velges i den adaptive kodebok vil imidlertid avvike fra en faktisk tonehøydeperiode i en tale, og det vil være umulig å fastslå om en tilstand av en inngangstale faktisk er støy eller ikke støy bare ut fra en verdi av tonehøydeperioden. Det problem at den kodede tale innenfor en støyfylt periode av talen er unaturlig kan da ikke løses på denne måte.

Foreliggende oppfinnelse har imidlertid som formål å løse de ovenfor angitte problemer. Særlig tar oppfinnelsen sikte på å frembringe fremgangsmåter for koding og dekoding av tale samt apparater for å gjengi tale med høy kvalitet selv ved lave bit-takter.

Beskrivelse av oppfinnelsen

Oppfinnelsen fremlegger en fremgangsmåte for taledekoding ifølge kodeeksitert lineærprediksjon (CELP), som kjennetegnes ved at den mottar en kodet tale og syntetiserer en tale ved å bruke minst en eksitasjonskodebok. Fremgangsmåten for taledekoding omfatter trinnene (1) å tilveiebringe en tidsserievektor med et antall sampler med null amplitudeverdi fra eksitasjonskodeboken, (2) å avgjøre om modifikasjon av tidsserievektoren er nødvendig, (3) å modifisere tidsserievektoren slik at antall sampler med null amplitudeverdi endres hvis modifikasjon avgjøres å være nødvendig, (4) en avgivelse av tidsserievektoren og (5) syntetisering av tale ved å benytte den avgitte tidsserievektoren.

Som et annet aspekt ved oppfinnelsen presenteres et apparat for taledekoding ifølge kodeeksitert lineær prediksjon (CELP), der taledekodingsanordningen mottar en kodet tale og syntetiserer en tale ved å bruke minst en eksitasjonskodebok. Apparatet kjennetegnes det omfatter en tidsserievektormoduleringsanordning for å tilveiebringe en tidsserievektor med et antall sampler med null amplitudeverdi fra eksitasjonskodeboken og å avgjøre om modifikasjon av tidsserievektoren er nødvendig, og for å modifisere tidsserievektoren slik at antall sampler med null

amplitudeverdi endres hvis modifikasjon avgjøres å være nødvendig, og for å avgi tidsserievektoren. Apparatet omfatter også en talesyntetiseringsanordning for å syntetisere en tale ved å bruke den avgitte tidsserievektoren.

5 Kort beskrivelse av tegningene

Fig. 1 viser et blokkskjema for en fullstendig konfigurasjon av et apparat for koding og dekoding av tale i en utførelse 1 i henhold til oppfinnelsen.

Fig. 2 viser en tabell for å forklare en evaluering av støynivået i den utførelse 1 i henhold til oppfinnelsen og som er vist i fig. 1.

10 Fig. 3 angir et blokkskjema for en fullstendig konfigurasjon av et apparat for koding og dekoding av tale i en utførelse 3 i henhold til foreliggende oppfinnelse.

Fig. 4 viser et blokkskjema av en fullstendig konfigurasjon av et apparat for koding og dekoding av tale i en utførelse 5 i henhold til oppfinnelsen.

15 Fig. 5 viser et skjematisk linjekart for å forklare en beslutningsprosess ved avveiling i en utførelse 5 som er vist i fig. 4.

Fig. 6 viser et blokkdiagram for en fullstendig konfigurasjon for et apparat for CELP-koding og -dekoding av tale i henhold til beslektet tidligere teknikk.

20 Fig. 7 viser et blokkskjema for en fullstendig konfigurasjon av et apparat for forbedret CELP-koding og -dekoding av tale i henhold til beslektet tidligere teknikk.

Beste modus for utførelse av oppfinnelsen

Utførelser av foreliggende oppfinnelse vil nå bli forklart med henvisning til tegningene.

25

Utførelse 1

Fig. 1 viser en fullstendig konfigurasjon som angir en fremgangsmåte for talekoding og taledekoding i en utførelse 1 i henhold til foreliggende oppfinnelse. I fig. 1 er det vist en koder 1, en dekode 2, en multiplekser 3 og en divisjonsenhet 4. Koderen 1 omfatter en analysator 5 for en lineær prediksjonsparameter, en koder 6 for lineær prediksjonsparameter, et syntesefilter 7, en adaptiv kodebok 8, en forsterkningskoder 10, en avstandsberegner 11, en første eksiteringskodebok 19, en andre eksiteringskodebok 20, en støynivåevaluator 24, en omkopler 25 for

eksiteringskodebøker og en innretning for avveiling og addering. Dekoderen 2 omfatter en dekoder 12 for lineær prediksjonsparameter, et syntesefilter 13, en adaptiv kodebok 14, en første eksiteringskodebok 22, en andre eksiteringskodebok 23, en støynivåevaluator 26, en omkopler 27 for eksiteringskodebøker, en forsterkingsdekoder 16, og en innretning 39 for avveiling og addering. I fig. 1 er analysatoren 5 for lineær prediksjonsparameter en spektruminformasjonsanalysator for å analysere en inngangstale S1 samt for å trekke ut en lineær prediksjonsparameter som angir spektruminformasjon for talen. Koderen 6 for prediksjonsparameter er en spektruminformasjonskoder for koding av den lineære prediksjonsparameter som gir spektruminformasjon og for å fastlegge en kodet lineær prediksjonsparameter som en koeffisient for syntesefilteret 7. De første eksiteringskodebøker 19 og 22 lagrer flere sett av støyfrie tidsserievektorer, og de andre eksiteringskodebøker 20 og 23 lagrer flere sett av tidsserievektorer med støy. Støynivåevaluatorene 24 og 26 evaluerer støynivåer, og omkoplerne 25 og 27 for omkopling av eksiteringskodebøker er innrettet for å kople om mellom slike kodebøker på grunnlag av støynivået.

Arbeidsoperasjonene vil nå bli forklart.

I koderen 1 analyserer analysatoren 5 for lineær prediksjonsparameter inngangstalen S1 og ekstraherer fra denne en lineær prediksjonsparameter som utgjør spektruminformasjon for vedkommende tale. Koderen 6 for lineær prediksjonsparameter koder denne lineære prediksjonsparameter. Koderen 6 for lineær prediksjonsparameter innstiller så en kodet lineær prediksjonsparameter som en koeffisient for syntesefilteret 7, og avgir også denne kodede lineære prediksjonsparameter til støynivåevaluatoren 24.

Kodingen av eksiteringsinformasjon vil nå bli forklart.

Tidligere eksiteringssignal er lagret i den adaptive kodebok 8, og en tidsserievektor som tilsvarende en adaptiv kode som er innført av avstandsberegningen 11, og som genereres ved å gjenta et tidligere eksiteringssignal periodisk, avgis på utgangssiden. Støynivåevaluatoren 24 evaluerer et støynivå i en bestemt kodingsperiode på grunnlag av den kodede lineære prediksjonsparameter som er tilført fra koderen 6 for lineær prediksjonsparameter og den adaptive kode, f.eks.

en spektrum-gradient, en korttids prediksjonsforsterkning og tonehøydefluktuasjon slik som vist i fig. 2, og avgir et evalueringsresultat til omkopleren 25 for eksitasjonskodebøker. Denne omkopler 25 for eksiteringskodebøker kopler om eksitasjonskodebøkene for koding basert på evalueringen av støynivået. Hvis f.eks. støynivået er lavt, benyttes den første eksiteringskodebok 19, og hvis støynivået er høyt anvendes den andre eksiteringskodebok 20.

Den første eksiteringskodebok 19 lagrer flere støyfrie tidsserievektorer, f.eks. flere tidsserievektorer som er opplært ved reduksjon av en forvrengning mellom en innlæringstale og dens kodede taleversjon. Den andre eksiteringskodebok 20 lagrer flere tidsserievektorer med støy, f.eks. et antall slike tidsserievektorer som er generert ut i fra tilfeldig støy. Såvel den første eksiteringskodebok 19 og den andre eksiteringskodebok 20 avgir en tidsserievektor som tilsvarer hver sin eksiteringskode som innføres fra avstandsberegneren 11. Hver av tidsserievektorene fra den adaptive kodebok 8 og enten den første eksiteringskodebok 19 eller den andre eksiteringskodebok 20 blir avveiet ved bruk av en tilsvarende forsterkning som avgis fra forsterkningskoderen 10 og adderes ved hjelp av innretningen 38 for avveilingen og addisjon. Et oppnådd addisjonsresultat avgis til syntesefilteret 7 som eksiteringssignaler, og det frembringes en kodet tale. Avstandsberegneren 11 beregneren en forskjell mellom den kodede tale og inngangstalen S1, samt søker en adaptiv kode, samt eksiteringskode og forstekning for å nedsette forskjellen til et minimum. Når kodingen er over avgis den lineære prediksjonsparameterkode og en adaptiv kode, eksiteringskode og forsterkningskode for nedsettelse av forvrengningsforskjellen mellom inngangstale og den kodede tale som et kodingsresultat S2. Dette er kjennetegnende arbeidsprosesser innenfor fremgangsmåten for talekoding i utførelse 1.

Dekoderen 2 vil nå bli forklart. I dekodeen 2 dekodeer dekodeenheten 12 for lineær prediksjonsparameter den lineære prediksjonsparameterkode til lineær prediksjonsparameter og innstiller denne kodede lineære prediksjonsparameter som en koeffisient for syntesefilteret 13, samt avgir på sin utgangsside den dekodete lineære prediksjonsparameter til støynivåevaluatoren 26.

Dekodingen av eksiteringsinformasjonen vil nå bli forklart. Den adaptive kodebok 14 avgir en tidsserievektor som tilsvarer en adaptiv kode og som genereres ved å gjenta et tidligere eksiteringssignal periodisk. Støynivåevaluatoren 26

evaluerer et støynivå ved å bruke den dekodete lineære prediksjonsparameter som er tilført fra dekodingsenheten 12 for lineær prediksjonsparameter og den adaptive kode på samme måte som støynivåevaluatoren 24 i koderen 1, og avgir et evalueringresultat til omkopleren 27 for eksiteringskodebøker. Denne omkopler 27 for eksiteringskodebøker kopler om mellom den første eksiteringskodebok 22 og den andre eksiteringskodebok 23 på grunnlag av evalueringresultatet med hensyn til støynivå på samme måte som eksiteringskodebok-omkopleren 25 i koderen 1.

Flere støyfrie tidsserievektorer, f.eks. flere slike vektorer som genereres ved innlæring for reduksjon av forvrengningsforskjell mellom en innlæringstale og dens tilsvarende kodede taleversjon, lagres i den første eksiteringskodebok 22. Flere tidsserievektorer med støy, f.eks. flere vektorer som er generert ut i fra tilfeldig støy, lagres i den andre eksiteringskodebok 23. Såvel den første som den andre eksiteringskodebok avgir en tidsserievektor tilsvarende hver sin eksiteringskode. Disse tidsserievektorer fra den adaptive kodebok 14 og enten den første eksiteringskodebok 22 eller den andre eksiteringskodebok 23 avveies ved anvendelse av forskjellige forsterkninger, som er dekodet fra forsterkningskodene fra forsterkningsdekoderen 16, samt adderes ved hjelp av innretningen 39 for avveining og addering. Et addisjonsresultat overføres til syntesefilteret 13 som et eksiteringssignal, og det frembringes da en utgangstale S3. Disse arbeidsoperasjoner er karakteristiske prosesser i utførelse 1 av fremgangsmåten for taledekoding.

I utførelse 1 evalueres støynivået i inngangstalen ved å bruke koden og kodingsresultatet, og forskjellige eksiteringskodebøker anvendes basert på evalueringresultatet. Tale av høy kvalitet kan således reproduseres ved hjelp av en forholdsvis liten datamengde.

I utførelse 1 er de flere tidsserievektorer lagret i hver av eksiteringskodebøkene 19, 20, 22 og 23. Denne utførelse kan virkeliggjøres i den utstrekning minst én tidsserievektor er lagret i hver av eksiteringskodebøkene.

30 Utførelse 2

I utførelse 1 utføres omkopling mellom to eksiteringskodebøker. Det er imidlertid også mulig å benytte tre eller flere eksiteringskodebøker som da omkoples i samsvar med et støynivå.

I utførelse 2 kan en egnet eksiteringskodebok anvendes til og med for en middelmådig tale, f.eks. med lett støy, i tillegg til de to nevnte taletyper, nemlig med og uten støy. På dette grunnlag kan således tale av høy kvalitet reproduseres.

6

Utførelse 3

Fig. 3 viser en fullstendig konfigurasjon for en fremgangsmåte for såvel talekoding som taledekoding i henhold til utførelse 3 for denne oppfinnelse. I fig. 3 anvendes samme henvisningstegn for komponenter som tilsvarer komponentene i fig. 1. I fig. 3 lagrer eksiteringskodebøkene 28 og 30 tidsserievektorer med støy, og prøvetakerne 29 og 31 innstiller en amplitudeverdi for en stikkprøve med lav amplitude til null i tidsserievektorene.

Arbeidsoperasjonene vil nå bli forklart. I koderen 1 analyserer analysatoren 5 for lineær prediksjonsparameter inngangstalen S1 og trekker ut fra denne en lineær prediksjonsparameter, som utgjør spektruminformasjon for vedkommende tale. Koderen 6 for lineær prediksjonsparameter koder så denne lineære prediksjonsparameter. Koderen 6 innstiller så den kodede lineære prediksjonsparameter som en koeffisient for syntesefilteret 7, og avgjør også på sin utgangsside den kodede lineære prediksjonsparameter til støynivåevaluatoren 24.

Kodingen av eksiteringsinformasjon vil nå bli forklart. Et tidligere eksiteringssignal er lagret i den adaptive kodebok 8, og en tidsserievektor som tilsvarer en adaptiv kode innført fra avstandsberegneren 11, og som genereres ved å gjenta det tidligere eksiteringssignal periodisk, avgis på utgangssiden. Støynivåevaluatoren 24 evaluerer et støynivå innenfor en tilsvarende kodingsperiode ved bruk av den kodede lineære prediksjonsparameter som er innført fra koderen 6 for lineær prediksjonsparameter, samt en adaptiv kode, f.eks. en spektrum-gradient, en korttids prediksjonsforsterkning samt tonehøydefluktuering, og avgir et evalueringresultat til prøvetakeren 29.

Eksiteringskodeboken 28 lagrer flere tidsserievektorer som f.eks. er generert fra tilfeldig støy, samt avgir en tidsserievektor som tilsvarer en eksiteringskode tilført fra avstandsberegneren 11. Hvis støynivået er lavt i evalueringresultatet for støyen, avgir prøvetakeren 29 en tidsserievektor hvor en amplitude for en stikkprøve med amplitudeverdi under en bestemt verdi innenfor de tidsserievektorer

som er tilført fra eksiteringskodeboken 28, f.eks. settes til null. Hvis støynivået er høyt, så avgir prøvetakeren 29 den tidsserievektor som er tilført fra eksiteringskodeboken 28 uten modifikasjon. Hver av tidsserievektorene fra den adaptive kodebok 8 og prøvetakeren 29 avveies ved anvendelse av en tilhørende forsterkning tilført fra forsterkningskoderen 10, samt adderes ved hjelp av innretningen 38 for avveiling og addisjon. Et addisjonsresultat avgis til syntesefilteret 7 som eksiteringssignaler, og kodet tale blir da frembrakt. Avstandsberegneren 11 beregner en forskjell mellom den kodede tale og inngangstalen S1, og søker etter en adaptiv kode, eksiteringskode og forsterkning for å nedsette denne forskjell til et minimum.

10 Når kodingen er ferdig, blir den lineære prediksjonsparameterkode og den adaptive kode, samt eksiteringskoden og forsterkningskoden for å nedsette forvrengningsforskjellen mellom inngangstale og den kodede tale avgitt som koderesultat S2. Disse prosesser utgjør da karakteristiske arbeidsoperasjoner i fremgangsmåten for talekoding i utførelse 3.

15 Dekoderen 2 vil nå bli nærmere forklart. I dekkeren 2 blir den lineære prediksjonsparameterkode dekodet til lineær prediksjonsparameter av dekkerenheten 12 for denne parameter. Denne dekkerenhet 12 for lineær prediksjonsparameter innstiller den lineære prediksjonsparameter som en koeffisient for syntesefilteret 13, og avgir også på sin utgangsside denne lineære prediksjonsparameter

20 til støynivåevaluatoren 26.

Dekodingen av eksitasjonsinformasjon skal nå bli forklart. Den adaptive kodebok 14 avgir en tidsserievektor som tilsvarer en adaptiv kode generert ved å gjenta et tidligere eksiteringssignal periodisk. Støynivåevaluatoren 26 evaluerer støynivå ved å anvende den dekodete lineære prediksjonsparameter som er inn-

25 ført fra dekkerenheten 12 for slik parameter samt den adaptive kode på samme måte som støynivåevaluatoren 24 i koderen 1, samt avgir et evalueringresultat til stikkprøveren 31.

Eksiteringskodeboken 30 avgir en tidsserievektor som tilsvarer en eksiteringskode. Stikkprøveren 31 avgir en tidsserievektor basert på evalueringresultatet med hensyn til støynivå ved samme prosess som utføres av stikkprøveren 29 i koderen 1. Hver av de tidsserievektorer som avgis fra den adaptive kodebok 14 og stikkprøveren 31 avveies ved å bruke en tilsvarende forsterkning som frembringes

av forsterkningskoderen 16, og vektorene adderes ved hjelp av innretningen 39 for avveining og addisjon. Et addisjonsresultat overføres til syntesefilteret 13 som et eksiteringssignal, og en utgangstale S3 blir da frembrakt.

I utførelse 3 opprettes en eksiteringskodebok som lagrer tidsserievektorer med støy, og en eksitering med et lavt støynivå kan genereres ved å ta eksiteringssignal-stikkprøver basert på et evalueringresultat av talens støynivå. Tale med høy kvalitet kan således reproduseres ut i fra en forholdsvis liten data-mengde. Da det således ikke er nødvendig å opprette flere eksiteringskodebøker, kan datalageret for å lagre eksiteringskodebøker reduseres.

10

Utførelse 4

I utførelse 3 blir tidsserievektorene enten punktprøvet eller ikke. Det er imidlertid også mulig å forandre en terskelverdi for en punktprøveamplitude basert på støynivået ved utprøvingen. I en utførelse 4 kan en egnet tidsserievektor genereres og også anvendes for en tale av midlere type, f.eks. én med svak støy, i tillegg til de to nevnte tale typer, nemlig med og uten støy. På denne måte kan således tale med høy kvalitet reproduseres.

Utførelse 5

Fig. 4 viser en fullstendig konfigurasjon for en fremgangsmåte for talekodning samt en fremgangsmåte for dekodning av tale i en utførelse 5 av foreliggende oppfinnelse, og samme henvisningstegn anvendes for enheter som tilsvarer enhetene i fig. 1.

I fig. 4 lagrer første eksiteringskodebøker 32 og 35 tidsserievektorer med støy, mens andre eksiteringskodebøker 33 og 36 lagrer støyfrie tidsserievektorer. Avveilingsbestemmere 34 og 37 er også vist.

Arbeidsfunksjonene vil nå bli forklart. I koderen 1 blir inngangstalen S1 analysert av analysatoren 5 for lineær prediksjonsparameter, og en slik lineær parameter ekstraheres og angir da spektruminformasjon for talen. Koderen 6 for lineær prediksjonsparameter vil da kode denne lineære prediksjonsparameter. Koderen 6 vil så innstille en kodet lineær prediksjonsparameter som en koeffisient for syntesefilteret 7, samt også avgi denne kodede prediksjonsparameter til støynivå-evaluatoren 24.

Kodingen av eksiteringsinformasjon vil nå bli forklart. Den adaptive kodebok 8 lagrer et tidligere eksiteringssignal og avgir en tidsserievektor som tilsvarer en adaptiv kode som er tilført fra avstandsberegneren 11, og som genereres ved å gjenta det tidligere eksiteringssignal periodisk. Støynivåevaluatoren 24 evaluerer et støynivå innenfor en tilsvarende kodingsperiode ved å anvende den kodede lineære prediksjonsparameter som er tilført fra koderen 6 for slik parameter samt den adaptive kode, f.eks. en spektrum-gradient, en kortsiktig prediksjonsforsterkning og en tonehøyde-fluktuasjon, og avgir et evalueringresultat til avveiiingsbestemmeren 34.

Den første eksiteringskodebok 32 lagrer flere tidsserievektorer med støy og som f.eks. er generert fra tilfeldig støy, og avgir en tidsserievektor som tilsvarer en eksiteringskode. Den andre eksiteringskodebok 33 lagrer flere tidsserievektorer som er generert ved innlæring og med henblikk på å redusere eventuell forvrengning mellom en innlæringstale og dens kodede taleuttrykk, samt for å avgi en tidsserievektor som tilsvarer en eksiteringskode tilført fra avstandsberegneren 11. Avveiiingsbestemmeren 34 fastlegger en avstemningsverdi som avgis til tidsserievektoren fra den første eksiteringskodebok 32 samt tidsserievektoren fra den andre eksiteringskodebok 33 basert på evalueringresultatet med hensyn til støynivå og som tilføres fra støynivåevaluatoren 24, slik som f.eks. angitt i fig. 5. Hver av tidsserievektorene fra den første eksiteringskodebok 32 og den andre eksiteringskodebok 33 avveies ved bruk av den vektverdi som avgis fra avveiiingsbestemmeren 34, og adderes til hverandre. I tidsserievektorer som avgis fra den adaptive kodebok 38 og de tidsserievektorer som genereres ved avveiiing og addering, blir avveiet ved bruk av tilsvarende forsterkninger som frembringes av forsterkningskoderen 10, samt adderes til hverandre ved hjelp av innretningen 38 for avveiiing og addisjon. Et addisjonsresultat avgis så til syntesefilteret 7 som eksiteringssignaler, og en kodetale blir da frembrakt. Avstandsberegneren 11 beregner en forskjell mellom den kodede tale og inngangstalen S1, og søker etter en adaptiv kode, eksiteringskode og forsterkning for å nedsette denne forskjell til et minimum. Når kodingen er fullført, blir så den lineære prediksjonsparameterkode, den adaptive kode, eksiteringskoden og forsterkningskoden for å nedsette forvrengningen mellom inngangstalen og den kodede tale, avgitt som et kodingsresultat.

Dekoderen 2 vil nå bli nærmere forklart. I dekode-
ren 2 dekodes den lineære prediksjonsparameterkode til lineær prediksjonsparameter ved hjelp av deko-
dingsenheten 12 for en slik lineær parameter. Denne dekodingsenhet 12 for lin-
eær prediksjonsparameter innstiller så den frembrakte lineære prediksjonspara-
meter som koeffisient for syntesefilteret 13, samt avgir også den lineære predik-
sjonsparameter til støyevaluatorene 26.

Dekodingen av eksiteringsinformasjon vil nå bli forklart. Den adaptive kode-
bok 14 avgir en tidsserievektor som tilsvarer en adaptiv kode ved å gjenta et tidlig-
ere eksiteringssignal periodisk. Støynivåevaluatorene 26 evaluerer et støynivå ved
å anvende den dekodete lineære prediksjonsparameter som er tilført fra dekod-
ingsenheten 12 for slik lineær parameter, samt den adaptive kode på samme
måte som ved støynivåevaluatorene 24 i koderen 1, samt avgir et evalueringssul-
tat til avveiiingsbestemmeren 37.

Den første eksiteringskodebok 35 og den andre eksiteringskodebok 36 av-
gir tidsvektorer som tilsvarer eksiteringskoder. Avveiiingsbestemmeren 37 foretar
avveiiing på grunnlag av det støynivå-evalueringssulatat som er tilført fra støynivå-
evaluatorene 26 etter samme metode som utføres av avveiiingsbestemmeren 34 i
koderen 1. Hver av tidsserievektorene fra den første eksiteringskodebok 35 og
den andre eksiteringskodebok 36 avveies ved bruk av en tilsvarende vektverdi
som frembringes av avveiiingsbestemmeren 37, og adderes til hverandre. Den
tidsserievektor som avgis fra den adaptive kodebok 14 og den tidsserievektor som
genereres ved avveiiing og addering, blir avveiet ved bruk av de tilsvarende for-
sterkninger dekodet ut i fra forsterkningskoder fra forsterkningsdekode-
ren 16, og addert til hverandre ved hjelp av innretningen 39 for avveiiing og addering. Det
frembringer således et addisjonsresultat som avgis til syntesefilteret 13 som et
eksiteringssignal, og en utgangstale S3 blir da frembrakt.

I utførelse 5 blir talens støynivå evaluert ved bruk av en kode og et kodings-
resultat, og tidsserievektoren med støy og den støyfrie tidsserievektor avveies på
grunnlag av evalueringssulatat, samt adderes til hverandre. Tale av høy kvalitet
kan da reproduseres ved hjelp av en forholdsvis liten datamengde.

Utførelse 6

I utførelsene 1-5 er det også mulig å forandre forsterkningskodebøker på grunnlag av evalueringresultatet med hensyn til støynivå. I utførelsen 6 kan det anvendes en meget egnet forsterkningskodebok basert på eksiteringskodeboken.
5 På denne måte kan da kvalitetstale reproduseres.

Utførelse 7

I utførelsene 1-6 blir talens støynivå evaluert og eksiteringskodebøkene omkoplet på grunnlag av evalueringresultatet. Det er imidlertid også mulig å fast-
10 legge og evaluere begynnende stemt tone, plosiv konsonant, etc., og utføre omkopling mellom eksiteringskodebøkene på grunnlag av et tilsvarende evalueringresultat. I utførelsen 7 er talen i tillegg til talens støytillstand klassifisert mer detaljert, f.eks. med hensyn til begynnende stemt tonetilstand, plosiv konsonant etc., og en egnet eksiteringskodebok kan da anvendes for hver tilstand. På denne måte
15 kan da høykvalitetstale reproduseres.

Utførelse 8

I utførelsene 1-6 blir støynivået innenfor kodingsperioden evaluert ved bruk av en spektrum-gradient, en kortsiktig prediksjonsforsterkning og tonehøydefluktu-
20 ering. Det er imidlertid også mulig å evaluere støynivået ved å bruke et forhold mellom en forsterkningsverdi og en utgangsverdi fra den adaptive kodebok.

Industriell anvendbarhet

Ved fremgangsmåtene for henholdsvis koding og dekoding av tale samt
25 apparatene for henholdsvis talekoding og taledekoding i henhold til foreliggende oppfinnelse blir støynivået for en tale evaluert innenfor en bestemt kodingsperiode ved å anvende en kode eller et kodingsresultat med hensyn til minst én av parametrene spektrum-informasjon, effektinformasjon og tonehøydeinformasjon, og forskjellige eksiteringskodebøker kan utnyttes på grunnlag av evalueringresultatet. På denne måte kan tale av høy kvalitet reproduseres av en forholdsvis liten
30 datamengde.

Ved fremgangsmåten for talekoding og taledekoding i henhold til foreliggende oppfinnelse er det opprettet flere eksiteringskodebøker som lagrer eksiteringer ved forskjellige støynivåer, og omkopling finner sted mellom disse flere eksiteringskodebøker basert på evaluering av talens støynivå. På denne måte kan høy kvalitetstale reproduseres ved bruk av en forholdsvis liten datamengde.

Ved fremgangsmåtene for henholdsvis talekoding og taledekoding i henhold til foreliggende oppfinnelse blir støynivåene for de tidsserievektorer som er lagret i eksiteringskodebøkene forandret på grunnlag av evaluering av talens støynivå. Ut i fra dette kan en høykvalitetstale reproduseres ved hjelp av en forholdsvis liten datamengde.

Ved fremgangsmåten for talekoding og fremgangsmåten for taledekoding i henhold til oppfinnelsen anvendes en eksiteringskodebok som lagrer tidsserievektorer med støy, og en tidsserievektor med lavt støynivå genereres ved punktprøving av signalstikkprøver i tidsserievektorene basert på evaluering av talens støynivå. Ut i fra dette kan tale med høy kvalitet reproduseres ved bruk av en forholdsvis liten datamengde.

Ved fremgangsmåter for henholdsvis talekoding og taledekoding i henhold til oppfinnelsen opprettes en første eksiteringskodebok, støy-tidsserievektorer og en annen eksiteringskodebok for lagring av støyfrie tidsserievektorer, og tidsserievektoren i den første eksiteringskodebok og tidsserievektoren i den andre eksiteringskodebok avveies på grunnlag av evalueringsresultatet med hensyn til støynivået i talen, samt addert til hverandre for å generere en tidsserievektor. En høykvalitetstale kan på denne måte reproduseres med bruk av forholdsvis liten datamengde.

PATENTKRAV

1. Fremgangsmåte for taledekoding ifølge kodeeksitert lineærprediksjon (CELP),

5 k a r a k t e r i s e r t v e d a t fremgangsmåten for taledekoding mottar en kodet tale og syntetiserer en tale ved å bruke minst en eksitasjonskodebok, idet fremgangsmåten for taledekoding omfatter:

å tilveiebringe en tidsserievektor med et antall sampler med null amplitudeverdi fra eksitasjonskodeboken;

10 å avgjøre om modifikasjon av tidsserievektoren er nødvendig;

hvis modifikasjon avgjøres å være nødvendig, å modifisere tidsserievektoren slik at antall sampler med null amplitudeverdi endres;

avgivelse av tidsserievektoren; og

syntetisering av en tale ved å benytte den avgitte tidsserievektoren.

15

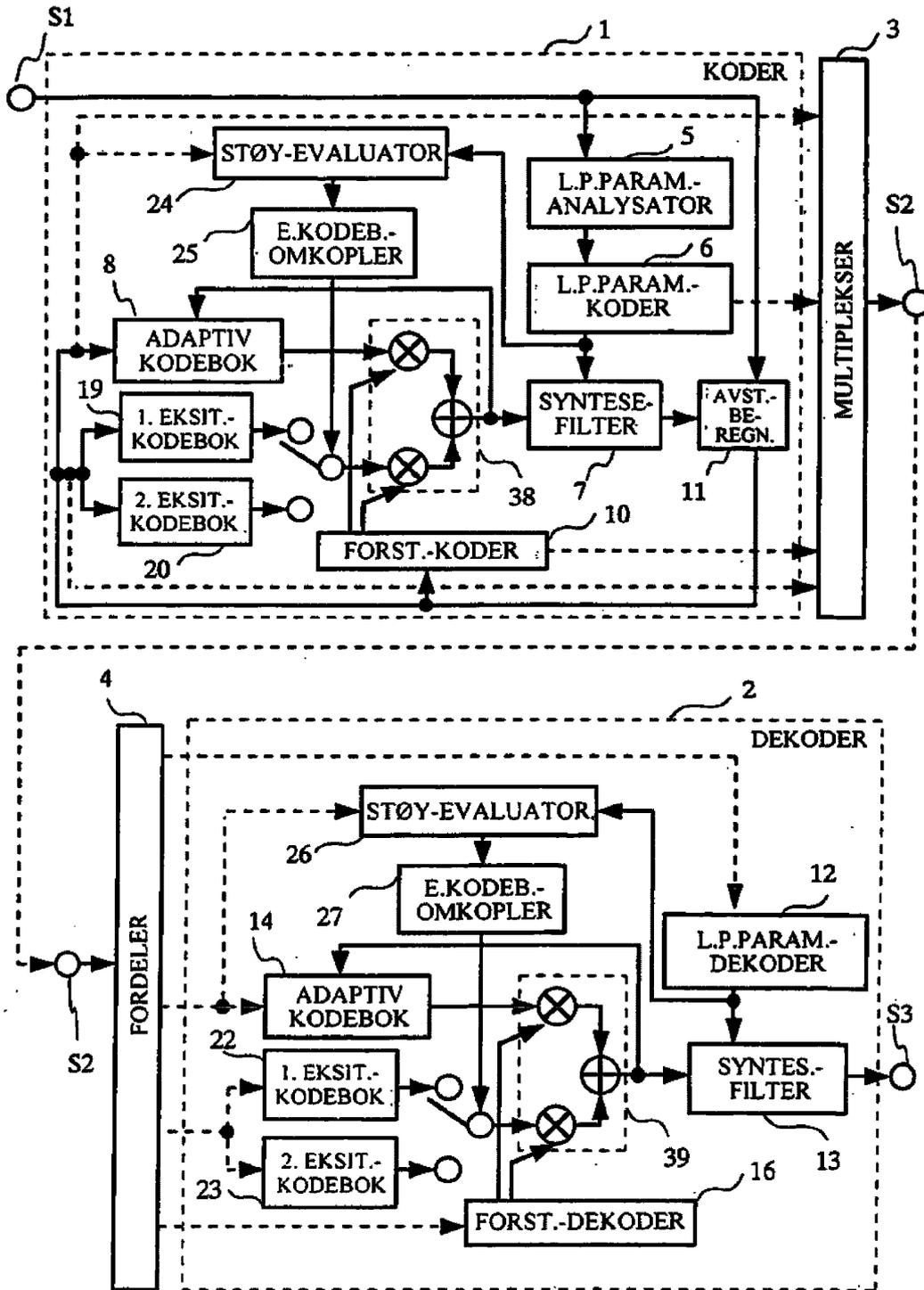
2. Apparat for taledekoding ifølge kodeeksitert lineær prediksjon (CELP), der taledekodingsanordningen mottar en kodet tale og syntetiserer en tale ved å bruke minst en eksitasjonskodebok

k a r a k t e r i s e r t v e d a t taledekodingsapparatet omfatter

20 en tidsserievektormoduleringsanordning for å tilveiebringe en tidsserievektor med et antall sampler med null amplitudeverdi fra eksitasjonskodeboken og å avgjøre om modifikasjon av tidsserievektoren er nødvendig, og for å modifisere tidsserievektoren slik at antall sampler med null amplitudeverdi endres hvis modifikasjon avgjøres å være nødvendig, og for å avgi tidsserievektoren; og

25 en talesyntetiseringsanordning for å syntetisere en tale ved å bruke den avgitte tidsserievektoren.

1/7
Fig.1

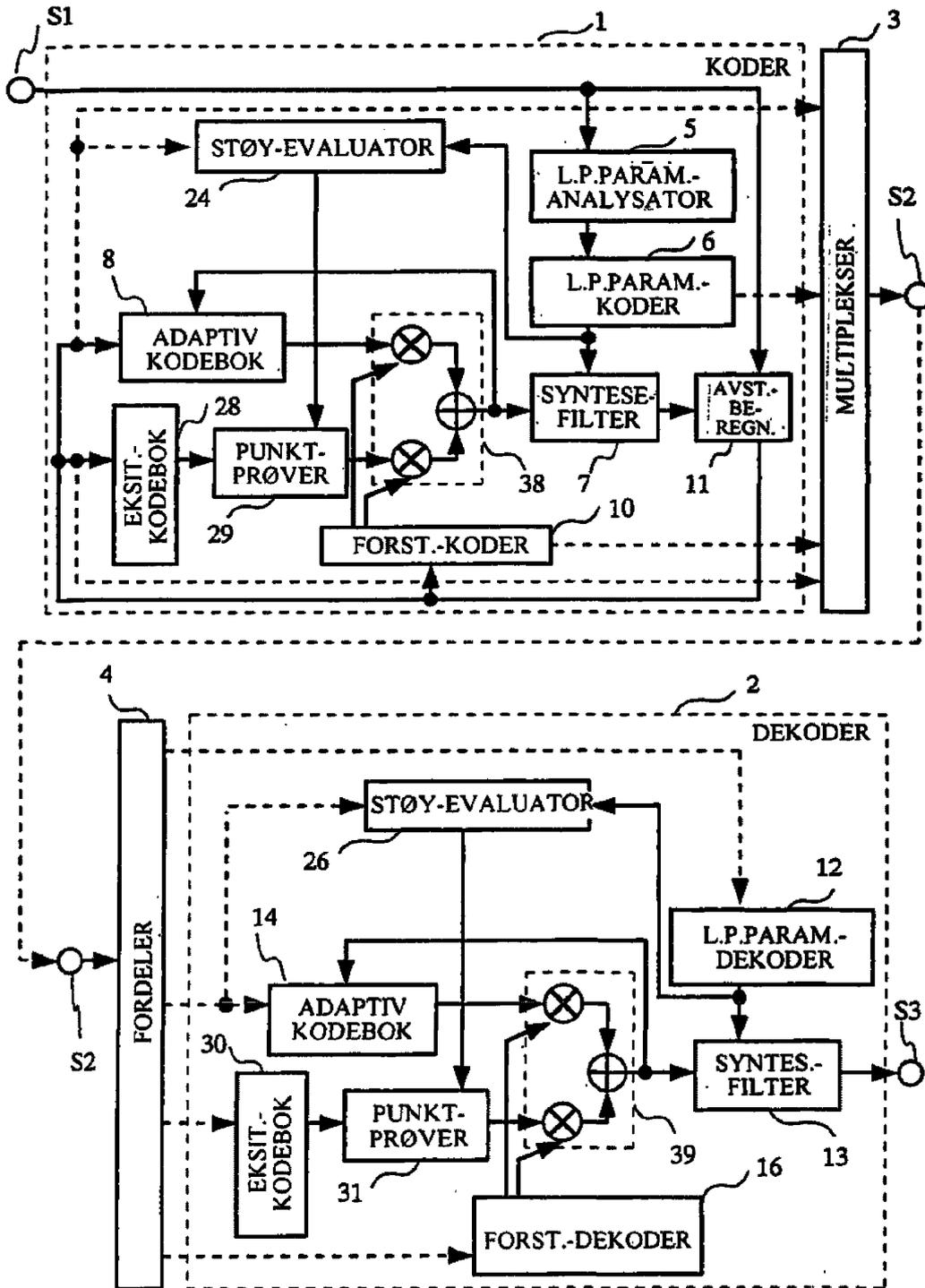


2/7

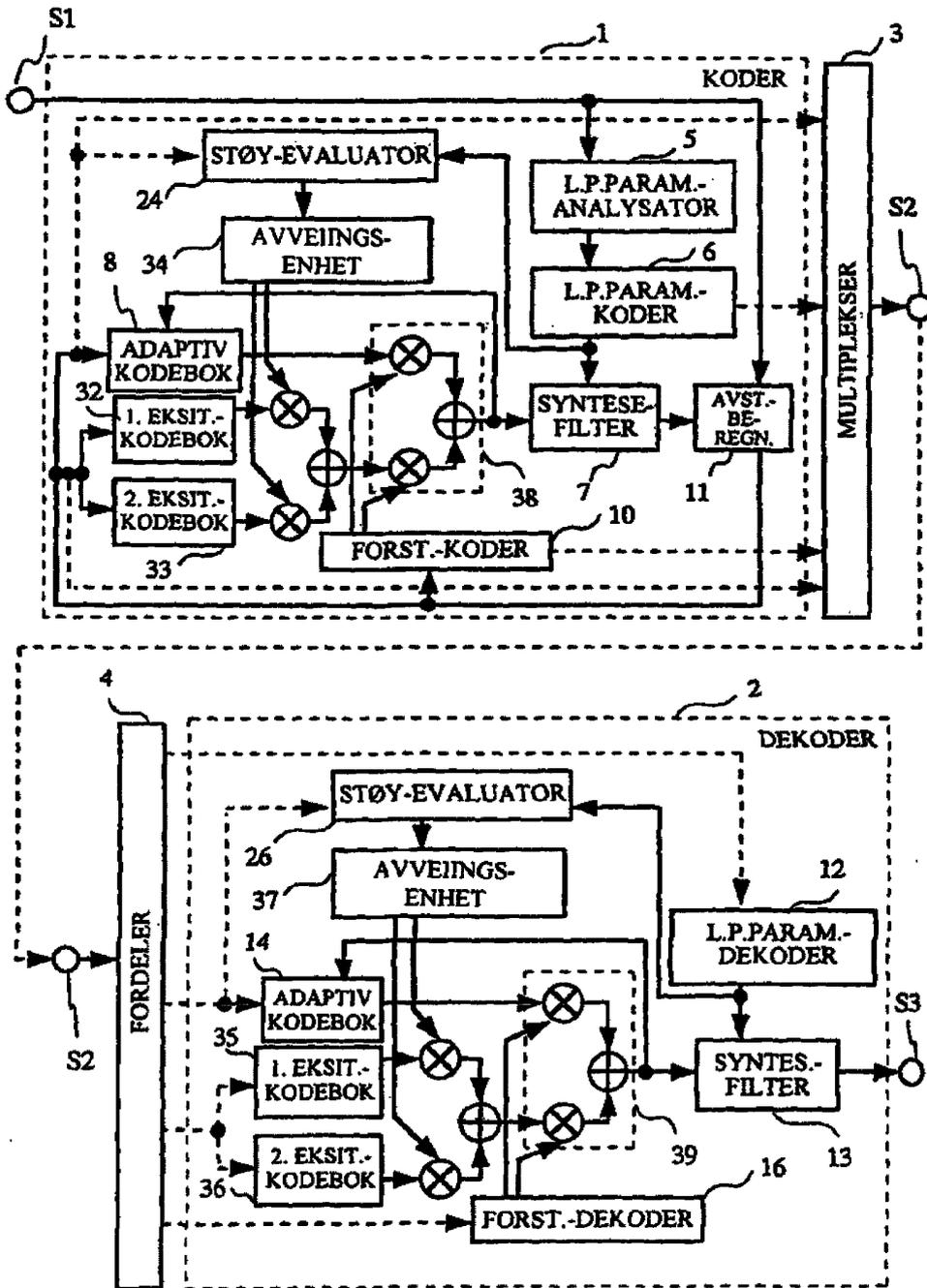
Fig.2

STØYNIVÅ	S ←————→ L
SPEKTRUM-GRADIENT	LAV GRADIENT ←————→ FLAT, HØY GRADIENT
KORTSIKTIG PREDIKSJONS-FORSTERKNING	L ←————→ S
TONEHØYDE-FLUKTUASJON	S ←————→ L

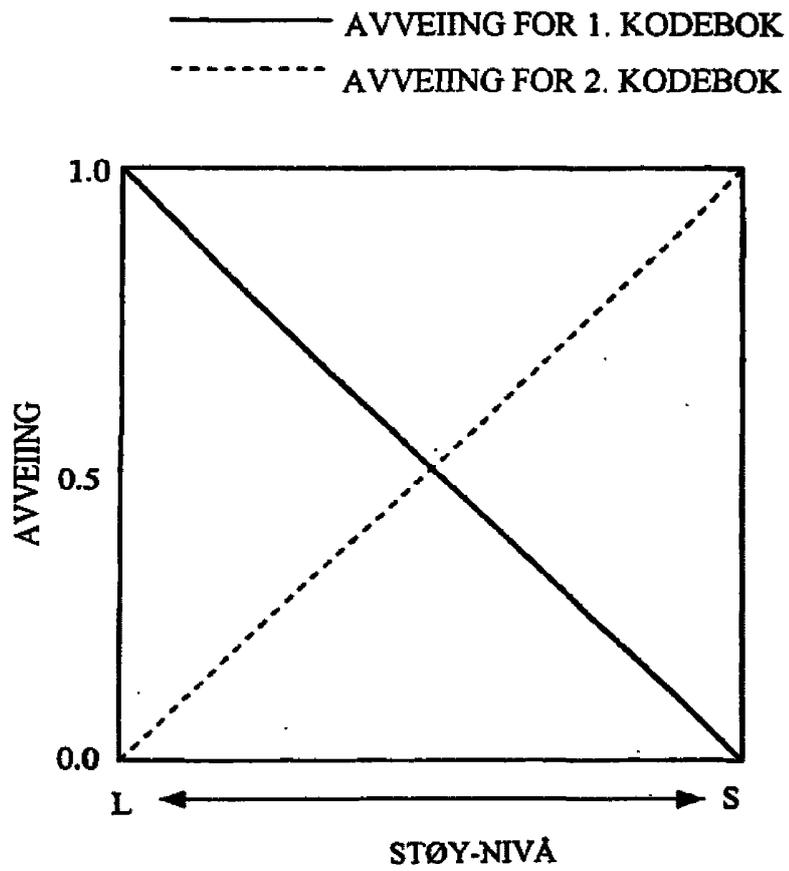
3/7
Fig.3



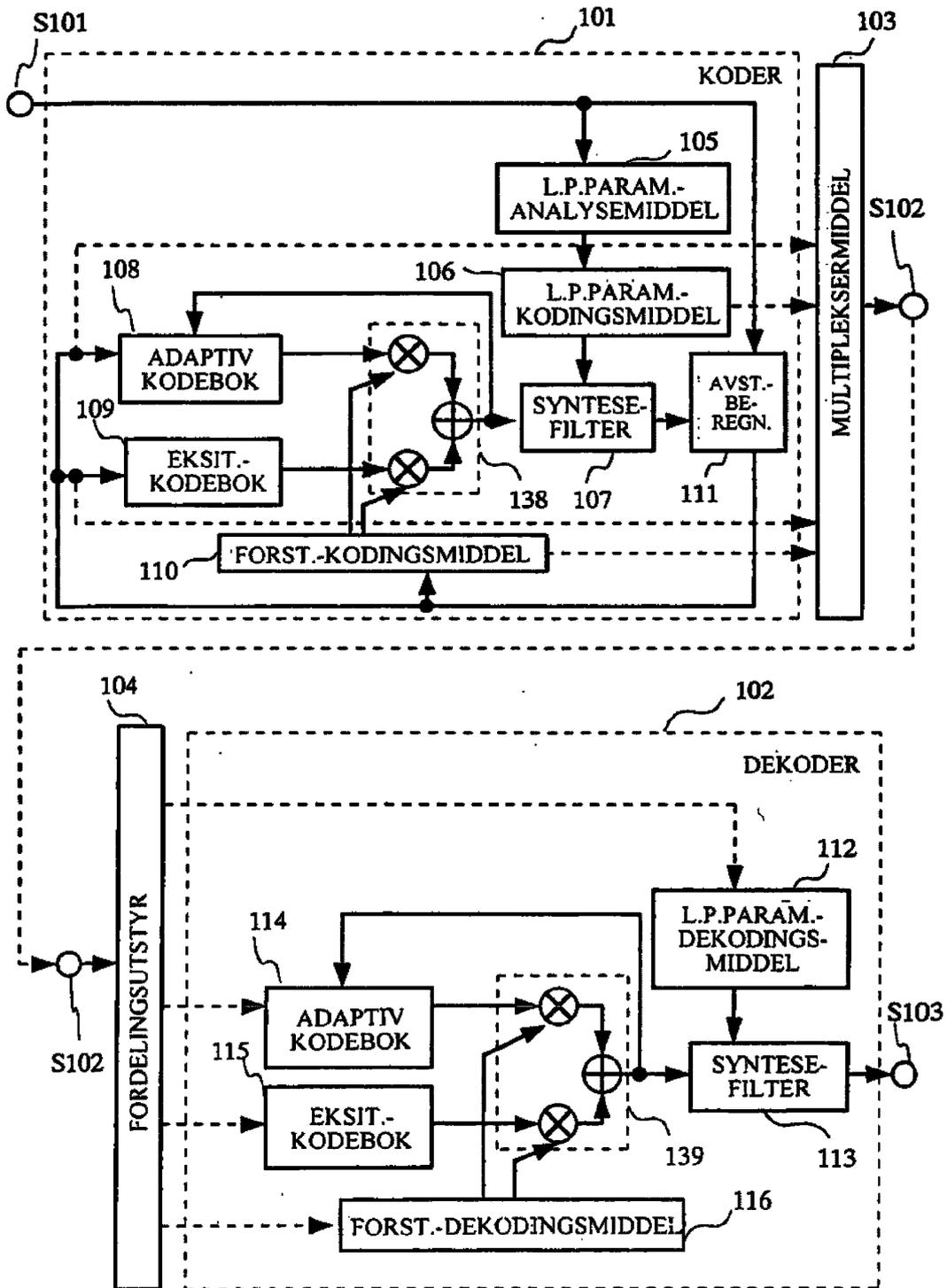
4/7
Fig.4



5/7
Fig.5



6/7
Fig.6



7/7
Fig.7

