

(19)



SUOMI - FINLAND

(FI)

PATENTTI- JA REKISTERIHALLITUS

PATENT- OCH REGISTERSTYRELSEN

FINNISH PATENT AND REGISTRATION OFFICE

(10) **FI 932465 A7**

(12) **JULKISEKSI TULLUT PATENTTIHAKEMUS
PATENTANSÖKAN SOM BLIVIT OFFENTLIG
PATENT APPLICATION MADE AVAILABLE TO THE
PUBLIC**

(21) Patentihakemus - Patentansökan - Patent application **932465**

(51) Kansainvälinen patenttiluokitus - Internationell patentklassifikation -
International patent classification (IPC⁷)
G10L 19/04

(22) Tekemispäivä - Ingningsdag - Filing date **28.05.1993**

(23) Saapumispäivä - Ankomstdag - Reception date **28.05.1993**

(41) Tullut julkiseksi - Blivit offentlig - Available to the public **02.12.1993**

(43) Julkaisupäivä - Publiceringsdag - Publication date **13.06.2019**

(32) (33) (31) Etuoikeus - Prioritet - Priority

01.06.1992 US 891596 25.06.1992 US 905992

(71) Hakija - Sökande - Applicant

1 • Hughes Electronics Corporation, N. Sepulveda Blvd. El Segundo, CA 90245-0956, USA, AMERIKAN YHDYSVALLAT, (US)

(72) Keksijä - Uppfinnare - Inventor

1 • Swaminathan, Kumar, USA, AMERIKAN YHDYSVALLAT, (US)

(74) Asiamies - Ombud - Agent

Heinänen Oy Patenttitoimisto, Airport Plaza, Äyritie 8 D, 01510 Vantaa

(54) Keksinnön nimitys - Uppfinningens benämning - Title of the invention

CELP-pohjainen puheenvaimennin

CELP-baserad talkompressor

CELP-PERUSTAINEN PUHEENVAIMENNIN

SELITYS

5 KEKSINNÖN TAUSTA

Keksinnön ala

Tämä keksintö liittyy yleisesti digitaalisiin äänikom-
munkaatiojärjestelmiin ja, erityisemmin, pienen bittinopeu-
den puhekoodekkiin, joka vaimentaa näytteitettyä puhetietoa
ja sitten vahvistaa vaimennetun puhetiedon takaisin alkupe-
räiseksi puheeksi. Sellaisia laitteita sanotaan yleisesti
"koodekeiksi" tarkoittaen kooderi/dekooderia. Keksinnöllä on
erityinen sovellutus digitaalisissa solukko- ja satelliitti-
kommunkaatioverkoissa, mutta sitä voidaan edullisesti
käyttää millä tahansa tuotelinjalla, joka vatii puheenvai-
mennusta telekommunkaatiota varten.

20 Tunnetun tekniikan kuvaus

Solukkotelekkommunkaatiojärjestelmät ovat kehittymässä
niiden nykyisestä analogisesta taajuusmoduloidusta (FM)
muodosta kohti digitaalijärjestelmiä. Telecommunication
Industry Association (TIA) on omaksunut standardin, joka
käyttää täyden nopeuden 8 Kbps:n vektorisummasyötettyä
lineaarisen ennustuksen (VSELP) puhekkooderia, konvolutionaa-
listaa koodausta virhesuojaukseen, differentiaalista 90
asteen vaihe-eron vaihesiirtoavainnuksen (QPSK) modulaatioi-
ta, ja aikajakoista, monipääsyistä (TDMA) kaaviota. Tämän on
odotettu kolminkertaistavan solukkojärjestelmien liikenteen-
kuljetuskapasiteetin. Jotta edelleen kasvatettaisiin sen
kapasiteettia kahdella, TIA on alkanut prosessin puolen
nopeuden koodekin arvioimiseksi ja seuraavaksi valitsemiseksi.
TIA:n teknologian arviointitarkoituksia varten puolen
nopeuden koodekilla pitäisi pitkin sen virhesuojausta olla
yleinen 6,4 Kbps:n bittinopeus ja olla rajoitettu 40 ms
kehyskokoon. Koodekilla on odotettu olevan täyden nopeuden

standardiin verrattava äänenlaatu laajalla olosuhteiden vaihtelualueella. Nämä olosuhteet käsittävät erilaiset kaiuttimet, käsiyksiköiden vaikutuksen, taustakohinaolosuhteet, ja kanavaolosuhteet.

5

Tehokas koodikirjasyötetty lineaarinen ennustus (Codebook Excited Linear Prediction) (CELP) tekniikka pienen nopeuden koodaukselle on nykyinen U.S. Federal standardi 4,8 Kbps:n CELP-kooderille. Kun CELP pitää suurimman lupauksen korkeasta äänen laadusta bittinopeuksilla 8,0 Kbps:n läheisyydessä, äänen laatu huononee 4 Kbps:ää lähestyvillä bittinopeuksilla. On tunnettua, että pääasiallisen lähteen laadun huononemiseksi on "ääneistetyt" puheen uudelleentuottamisessa. CELP-kooderin perustekniikka käsittää satunnaisesti jaettujen syöttövektoreiden koodikirjan etsimisen sille vektorille, joka tuottaa ulostulosekvenssin (kun suodatettu virityksen ja lineaarisesti ennustavan koodauksen (LPC) lyhyen ajan synteesisuotimien kautta), joka on lähinnä sisääntulosekvenssiä. Tämän tehtävän suorittamiseksi kaikki ehdolla olevat syöttövektorit koodikirjassa täytyy olla suodatettu sekä viritys- että LPC-synteesisuotimien kanssa, jotta tuotetaan ehdolla oleva ulostulosekvenssi, jota voidaan siten verratta sisääntulosekvenssi. Tämä tekee CELP:stä hyvin laskennallisesti intensiivisen algoritmin, jolla on tyypillisesti koodikirjoja käsittäen 1024 sisääntuloa tai enemmän. Lisäksi käytetään havaintokykyistä virheen punnitsevaa suodinta, joka lisää laskentakuormaa. Nopeita digitaalisia signaaliprosessoreita on autettu suorittamaan hyvin monimutkaisia algoritmeja, kuten CELP, reaaliajassa, mutta korkean äänen laadun kehittämisen ongelma pienillä bittinopeuksilla jää. Jotta saataisiin koodekkeja telekommunikaatiolaitteistoihin, äänen laadun tarvitsee olla verrattavissa 8,0 Kbps:n digitaaliseen solukkostandardiin.

20

25

30

35

KEKSINNÖN YHTEENVETO

Tämä keksintö antaa tekniikan korkealaatuiselle pienen bittinopeuden puhekoodekille, joka käyttää parannettua CELP-syöttöanalyysiä ääneistetylle puheelle, joka voi kehittää äänen laadun, joka on verrattavissa täyden nopeuden koodekin laatuun, jota käytetään North American Digital Cellular standardissa ja soveltuu siksi käytettäväksi telekommunikaatiolaitteistossa. Keksintö antaa telekommunikaatioluokkaisen koodekin, joka lisää solukkokanavan kapasiteettia kahdella.

Eräässä tämän keksinnön edullisessa sovellutusmuodossa pienen nopeuden koodekki, joka käyttää ääneistettyä puheensyöttömallia, vaimentaa mitä tahansa puhetietoa näytteistetyinä 8 KHz:lla, esim.,. 64 Kbps:n PCM, 4,2 Kbps:ään ja vahvistaa sen takaisin alkuperäiseksi puheeksi. Siihen liittyvä huononnus äänenlaadussa on verrattavissa IS54-standardin 8,0 Kbps:n äänikooderiin, jota käytetään U.S.:n digitaalisissa solukkojärjestelmissä. Tämä on suoritettu käyttämällä samaa parametristä mallia, jota käytetään traditionaalisissa CELP-koodereissa, mutta määrittämällä ja päivittämällä nämä parametrit erikseen kahdessa eri moodissa (A ja B) vastaten stationäärisiin ääneistettyihin puhesegmentteihin ja ei-stationäärisiin ääneistämättömiin puhesegmentteihin. Pienen nopeuden puhedekooderi on kuten useimmat CELP-dekooderit lukuunottamatta sitä, että se toimii kahdessa moodissa riippuen vastaanotetusta moodibitistä. Sekä virityksen esisuodatusta että yleistä jälkisuodatusta käytetään synteisoidun puheen korottamiseksi.

Edellä mainitun keksinnön erityisen sovellutusmuodon mukainen Pienen bittinopeuden koodekki käyttää 40 ms puhekehyksiä. Kussakin puhekehyksessä puolen nopeuden puhekoooderi suorittaa LPC-analyysiä kahdessa 30 ms nopeusikkunassa, jotak on erotettu erilleen 20 ms:llä. Ensimmäinen ikkuna on keskitetty keskelle, ja toinen ikkuna on keskitetty 40 ms puhekehyn reunaan. Kaksi virityksen estimaattia on määritetty käyttäen puheikkunoita, jotka, kuten LPC-analyysi-ik-

kunat, on keskitetty keskelle ja 40 ms puhekehyyksen reunaan. Virityksen estimointialgoritmi sisältää sekä taaksepäin että eteenpäin menevän virityksen seurannan ensimmäiselle virityksenanalysointi-ikkunalle, mutta vain taaksepäin menevän virityksen seurannan toiselle virityksen analysointi-ikkunalle.

Perustuen kahteen silmukkavirityksen estimaattiin ja kahteen kvantisoitujen suodinkertoimien yksikköön puhekehys on luokiteltu kahteen moodiin. Yksi moodi on pääasiassa ääneistetty ja on tunnettu hitaasti muuttuvasta äänialan muodosta ja hitaasti muuttuvasta äänisoinnun värähdysnopeudesta tai -virityksestä. Tämä moodi on nimetty moodi A:ksi. Toinen moodi on etupäässä ääneistämätön ja on nimetty moodi B:ksi. Moodi A:ssa toinen virityksen estimaatti on kvantisoitu ja lähetetty. Tätä on käytetty ohjaamaan suljetun silmukan virityksen estimointia kussakin alikehyksessä. Moodivalinnan kriteeri käyttää kahden virityksen estimaattia, kvantisoituja suodinkertoimia toiselle LPC-analyysi-ikkunalle, ja kvantisoimattomia suodinkertoimia ensimmäiselle LPC-analyysi-ikkunalle.

Tämän keksinnön eräässä edullisessa sovellutusmuodossa moodille A 40 ms:n puhekehys on jaettu seitsemään alikehykseen. Kuusi ensimmäistä ovat pituudeltaan 5,75 ms j seitsemäs on pituudeltaan 5,5 ms. Kussakin alikehyksessä viritysindeksi, virityksen vahvistuksen indeksi, kiinteän koodikirjan indeksi, kiinteän koodikirjan vahvistusindeksi ja kiinteän koodikirjan vahvistuksen merkki on määrätty käyttämällä synteessin lähestymisen analyysiä. Suljetun silmukan virityksen indeksin hakualue on keskitetty kvantisoidun viritysestimaatin ympärille, joka on saatu nykyisen 40 ms:n kehyksen toisesta viritysanalyysi-ikkunasta yhtä hyvin kuin edellisen 40 ms:n kehyksen vastaavasta, jos se oli moodin A kehys tai aikaisemman 40 ms:n kehyksen viimeisen alikehyksen viritys, jos se oli moodin B kehys. Suljetun silmukan viritysindeksin hakualue on 6-bittinen hakualue kussakin alike-

hyksessä, ja se käsittää sekä murtoluku- että kokonaisluku-
viritysviiveitä. Suljetun silmukan viritysvahvistus on
kvantisoitu hakusilmukan ulkopuolella käyttämällä kolmea
bittiä kussakin alikehyksessä. Virityksen vahvistuksen
5 kvantisointitaulukot ovat erilaisia kummassakin moodissa.
Kiinteä koodikirja on 6-bittinen glottaalisen pulssin koodi-
kirja, jonka vierekkäisillä vektoreilla on kaikki paitsi sen
loppuelementit yhteisiä. Käytetään hakuproseduuria, joka
käyttää hyväksi tätä. Yhdessä tämän keksinnön edullisessa
10 sovellutusmuodossa kiinteän koodikirjan vahvistus on kvan-
tisoitu käyttämällä neljää bittiä alikehyksissä 1, 3, 5 ja 7
ja käyttämällä rajoitettua 3-bitin aluetta keskitettynä
aikaisemman alikehysvahvistuksen indeksin ympärille alike-
hyksille 2, 4 ja 6. Sellainen differentiaalinen vahvistuksen
15 kvantisointikaavio ei ole vain tehokas käytettyjen bittien
suhteen, vaan myös vähentää kiinteän koodikirjan haku-
proseduurin monimutkaisuutta, koska vahvistuksen kvan-
tisointi on tehty hakusilmukassa. Lopulta kaikki yllä maini-
tut parametriesimaatit on puhdistettu käyttämällä viivästet-
20 tyä päätöslähestymistä. Siten, kussakin alikehyksessä,
suljetun silmukan virityshaku tuottaa M parasta estimaattia.
Kullekin näistä M parhaalle viritysestimaatille ja N par-
haalle aikaisemmalle alikehysparametrille, saadaan MN opti-
maalista viritysvahvistusindeksiä, Kiinteän koodikirjan
25 indeksit, kiinteän koodikirjan vahvistusindeksit, ja kiinte-
än koodikirjan vahvistusindeksit. alikehyksen lopussa nämä
MN ratkaisua on karsittu L parhaaseen käyttämällä kumulatii-
vissa signaali-kohinasuhdetta (SNR) kriteerinä. Ensimmäisel-
le alikehykselle, $M=2$, $N=1$, $L=2$ on käytetty. Viimeiselle
30 alikehykselle $M=2$, $N=2$, $L=1$ on käytetty, kun taas toisille
alikehyksille $M=2$, $N=2$, $L=2$ on käytetty. Viivästetty päätös-
lähestyminen on erityisesti tehokas ääneistetyin siirrossa
ääneistämättömään ja ääneistämättömän siirrossa ääneistetty-
jen alueille. Lisäksi, se johtaa tasaisempaan viritysrataan
ääneistetyllä alueella. Tämä viivästetty päätöslähestyminen
35 johtaa N kertaa suljetun silmukan virityshaun monimutakisuus-
teen, mutta paljon vähemmän kuin MN kertaa kiinteän koodi-

kirjan haun monimutkaisuuteen kussakin alikehyksessä. Näin on, koska vain korrelaatiotermi tarvitsee laskea MN kertaa kiinteälle koodikirjalle kussakin alikehyksessä, mutta energiatermi tarvitsee laskea vain kerran.

5

Moodille B 40 ms:n puhekehys on jaettu viiteen alikehykseen, joiden kunkin pituus on 8 ms. Kussakin alikehyksessä viritysindeksi, viritysvahvistuksen indeksi, kiinteän koodikirjan indeksi, ja kiinteän koodikirjan vahvistuksen indeksi on määritetty käyttämällä suljetun silmukan analyysia synteesillä lähestymisellä. Suljetun silmukan viritysindeksihaun alue ulottuu koko alueelle 20:sta 146:een. Vain kokonaislukuviritysviiveitä käytetään. Avoimen silmukan viritysestimaatit on jätetty huomioon ottamatta eikä käyetty tässä moodissa. Suljetun silmukan viritysvahvistus on kvantisoituhakusilmukan ulkopuolella käyttämällä kolmea bittiä kussakin alikehyksessä. Viritysvahvistuksen kvantisointitaulukot ovat erilaisia kahdessa moodissa. Kiinteä koodikirja on 9-bittinen moni-innovaatiokoodikirja käsittäen kaksi osaa. Yksi on Hadamar-vektorisummaosa ja toinen on sinkkipulssiosa. Tämä koodikirja käyttää hakuproseduuria, joka käyttää hyväksi näiden osien rakennetta ja takaa positiivisen vahvistuksen. Kiinteän koodikirjan vahvistus on kvantisoitu käyttämällä neljää bittiä kaikissa alikehyksissä hakusilmukan ulkopuolella. Kuten huomautettiin aikaisemmin, vahvistuksen on taattu olevan positiivinen ja siksi ei merkkibittiä tarvitse lähettää kunkin kiinteän koodikirjan vahvistusindeksin kanssa. Lopulta, kaikki yllä olevista parametriestimaateista on puhdistettu käyttämällä viivästettyä päätöslähestymistä, joka on identtinen moodi A:ssa käytettyyn.

30

LYHYT PIIRUSTUSTEN KUVAUS

Edellä esitetty ja muut kohteet, aspektit ja edut ymmärrettään paremmin seuraavasta keksinnön edullisen sovellutusmuodon yksityiskohtaisesta kuvauksesta viittauksella piirustuksiin, joissa:

35

KUV. 1 on lähettimen lohkokaavio johdottomassa kommunikaatiojärjestelmässä, joka käyttää pienen bittinopeuden puhekoodausta keksinnön mukaisesti;

- 5 KUV. 2 on vastaanottimen lohkokaavio johdottomassa kommunikaatiojärjestelmässä, joka käyttää pienen bittinopeuden puhekoodausta keksinnön mukaisesti;

- 10 KUV. 3 on lohkokaavio lähettimestä, jota käytetään KUVIOssa 1 esitetyssä lähettimessä;

KUV. 4 on dekooderin lohkokaavio, jota käytetään KUV. 2 esitetyssä vastaanottimessa;

- 15 KUV. 5A on ajoitusdiagrammi esittäen lineaarisen ennustusanalyysi-ikkunoiden poikkeaman keksinnön toteutuksessa;

- 20 KUV. 5B on ajoitusdiagramma esittäen virityksen ennustusanalyysi-ikkunoiden poikkeaman avoimen silmukan viritysennustukselle keksinnön toteutuksessa;

KUV. 6 on vuokaavio valaisten keksinnön 26-bittisen linjan spektritaajuusvektorin kvantisointiprosessin;

- 25 KUV. 7 on vuokaavio valaisten tunnetun viritysseurantaalgoritmin toimintaa;

- 30 KUV. 8 on lohkokaavio esittäen yksityiskohtaisemmin KUV. 3 esitetyn kooderin avoimen silmukan viritysestimoinnin toteutuksen;

KUV. 9 on vuokaavio valaisten modifioidun viritysseurantaalgoritmin toimintaa toteutettuna KUV. 8 esitetyllä avoimen silmukan viritysestimoinnilla;

- 35 KUV. 10 on lohkokaavio esittäen yksityiskohtaisemmin KUV. 3 esitetyn kooderin moodin määrittelyn toteutuksen;

KUV. 11 on vuokaaviuo valaisten moodivalintaproseduuria toteutettuna KUV. 10 esitetyllä moodinmäärittäyspiirillä;

5 KUV. 12 on ajoitusdiagramma esittäen alikehyksen rakenteen moodissa A;

KUV. 13 on lohkokaaavio esittäen yksityiskohtaisemmin KUV. 3 esitetyn kooderin syöttömallinnuspiirin toteutuksen;

10 KUV. 14 on käyrä esittäen glottaalisen pulssimuodon;

KUV. 15 on ajoitusdiagramma esittäen jäljitysvarmistuksen esimerkin viivästetyn päätöksen jälkeen moodissa A; ja

15 KUV. 16 on lohkokaaavio esittäen keksinnön mukaisen puhedekooderin toteutuksen.

KEKSINNÖN EDULLISEN SOVELLUTUSMUODON YKSITYISKOHTAINEN KUVAUS

20

Viitaten nyt piirustuksiin, ja erityisemmin KUV. 1, on esitetty lohkokaaavio lähettimestä johdottomassa kommunikatiojärjestelmässä, joka käyttää keksinnön mukaista pienen bittinopeuden puhekoodausta. Analoginen puhe, sopivasta käsiyksiköstä, on näytteistetty 8 kHz:n nopeudella ja muunnettu digitaalisiksi arvoiksi analogia/digitaali(A/D)-muuntimella 11 ja syötetty puhekooderiin 12, joka on tämän keksinnön kohteena. Koodattu puhe on edelleen koodattu kanavakooderilla 13, kuten voi olla vaadittu, esimerkiksi, digitaalisessa solukkokommunikatiojärjestelmässä, ja lopputuloksena saatava koodattu bittivirta on syötetty modulaattoriin 14. Tyypillisesti, vaihesiirtoavainnusta (PSK) on käytetty ja, siksi, modulaattorin 14 ulostulo on muunnettu digitaali/analogia(D/A)-muuntimella 15 PSK-signaaleiksi, jotka on vahvistettu ja taajuuskerrottu radiotaajuisella (RF) ylösmuuntimella 16 ja säteilee antennilla 17.

- 5 Analogisen puhesignaalin sisääntulon järjestelmään on oletettu olevan alipäästösuodatettu käyttämällä ei-poikkeavaa suodinta ja näytteistetty 8 kHz:llä. Digitoidut näytteet A/D-muuntimesta 11 on ylipäästösuodatettu ennen mitään prosessointia käyttämällä toisen asteen neliasteista suodinta, jolla on siirtofunktiio

$$H_{HP}(Z) = \frac{(1 - 2Z^{-1} + Z^{-2})}{1 - 1.8891Z^{-1} + 0.89503Z^{-2}}$$

- 10 Ylipäästösuodatinta käytetään pienentämään kaikkea tasavirtaa tai huminapitoisuutta sisääntulevassa puhesignaalissa.

- 15 KUV. 2 lähetetty signaali on vastaanotettu antennilla 21 ja sekoitettu välitystaajuuteen (IF) RF-alasmuuntimella 22. IF-signaali on muunnettu digitaaliseksi biitivirraksi A/D-muuntimella 23, ja lopputuloksena saatava biitivirta on tulkittu demodulaattorissa 24. Tässä kohdassa tapahtuu koodausprosessille käänteinen lähettimessä. Erikoisesti, tulkinta on suoritettu kanavadekooderilla 25 ja puhedekooderilla 26, joista jälkimmäinen on myös tämän keksinnön kohteena. Lopulta puhedekooderin ulostulo on syötetty D/A-konvertteriin 27, jossa on 8 kHz:n näytteenottotaajuus analogisen puhen syntetisoimiseksi.

- 20
25 KUV. 1 kooderi 12 on esitetty KUV. 3 ja sisältää ääniesiprosessorin 31, jota seuraa lineaarinen ennustus(LP) analyysi ja kvantisointi lohkoissa 32. Perustuen lohkon 32 ulostuloon, viritysestimointi on tehty lohkoissa 33 ja moodin määrittäminen moodissa A tai moodissa B kuten kuvattu yksityiskohtaisemmin jälkeempinä, on tehty lohkoissa 34. Moodi, kuten määritetty lohkoissa 34, määrittää syöttömallinnuksen lohkoissa 35, ja tätä seuraa vaimennettun puheen bittien pakkaaminen prosessorilla 36.

KUV. 2 dekooderi on esitetty KUV. 4 ja sisältää prosessorin 41 vaimennettujen puhebittien avaamiseksi. Avatut puhebitit on käytetty lohossa 42 syöttösignaalin uudellenkonstruoimiseksi, mitä seuraa virityksen esisuodatus suodattimessa 43.

5 Suodattimen 43 ulostulo on edelleen suodatettu puhesyntaasisuodattimessa 44 ja globaalissa jälkisuodattimessa 45.

KUV. 3 Pienen bittinopeuden koodekki käyttää 40 ms:n puhekehys-
 10 hyksiä. Kussakin puhekehyksessä pienen nopeuden puhekooderi suorittaa LP(lineaarinen ennustus) analyysiä lohossa 32 kahdessa 30 ms:n puheikkunassa, jotak on erotettu toisistaan 20 ms:llä. Ensimmäinen ikkuna on keskitetty keskelle ja toinen ikkuna on keskitetty 40 ms:n puhekehyn loppuun.

15 Molempien LP-analyysi-ikkunoiden poikkeama on esitetty KUV. 5A. Kukin LP-analyysi-ikkuna on kerrottu Hamming-ikkunalla ja sitä seuraa LP-analyysin kymmenennen asteen autokorrelaatiomenetelmällä. Molempien suodinkertoimien yksiköt on kaistaleveydeltään levennetty 15 Hz:lla ja munnettu linjaspektritaajuuksiksi. Nämä kymmenen linjaspektritaajuutta on kvantisoitu 26-bittisellä LSF VQ:lla tässä sovellutusmuodossa. Tätä 26-bittistä LSF VQ:ta on kuvattu seuraavaksi.

20

25
30
35

Kymmenen linjaspektritaajuutta molemmille yksiköille on kvantisoitu lohossa 32 26-bittisellä monikoodikirjan jaettulla vektorikvantisoijalla. Tämä 26-bittinen LSF-vektori-
 kvantisoija luokittelee kvantisoimattoman linjaspektritaajuuden vektorin "ääniIRS-suodatettuna", "ääneistämättömänä IRS-suodatettuna", "ääneistettynä ei-IRS-suodatettuna", ja "ääneistämättömänä ei-IRS-suodatettuna" vektorina, jossa "IRS" viittaa välireferenssijärjestelmäsuodattimeen, kuten on spesifioitu CCITT:ssä, Blue Book, Rec. P.48. LSF-vektorin kvantisointiprosessin esitys on esitetty KUV. 6 vuokaavion muodossa. Kullekin luokityukselle jaettu vektroikvantisoijaa on käytetty. "Ääneistetylle IRS-suodatetulle" ja "ääneistetylle ei-IRS-suodatetulle" kategorioille 51 ja 53, 3-4-3 jaettua vektroikvantisoijaa on käytetty. Ensimmä-

mäiset kolme LSFää käyttää 8-bittistä koodikirjaa toimintalo-
 lohkoissa 55 ja 57, seuraavat neljä LSFää käyttää 10-bittis-
 tä koodikirjaa toimintalohkoissa 59 ja 61, ja viimeiset
 kolme LSFää käyttävät 6-bittistä koodikirjaa toimintaloh-
 koissa 63 ja 65. "Ei-ääneistetyille IRS-suodatetulle" ja
 "ääneistämättömälle ei-IRS-suodatetulle" kategorioille 52 ja
 54 3-3-4 jaettu vektorikvantisoijaa on käytetty. Ensimmäiset
 kolme LSFää käyttävät 7-bittistä koodikirjaa toimintaloh-
 koissa 56 ja 58, seuraavat kolme LSFää käyttävät 8-bittistä
 vektorikoodikirjaa toimintalohkoissa 60 ja 62, ja viimeiset
 neljä LSFää käyttävät 9-bittistä koodikirjaa toimintaloh-
 koissa 64 ja 66. Kustakin jaetun vektroin koodikirjasta
 kolme parasta ehdokasta on valittu toimintalohkoissa 67, 68,
 69 ja 70 käyttämällä energiapainotettua keskimääräistä
 neliöjuurivirheen kriteeriä. Energiapainotus heijastaa
 spektriverhokäyrän tehotasoa kullakin linjaspektritajuudel-
 la. Kolme parasta ehdokasta kullekin kolmesta jaetusta
 vektorista johtaa yhteensä kahteenkymmeneen seitsemään
 yhdistelmään kullekin kategorialle. Haku on pakotettu, niin
 että ainakin yksi yhdistelmä päätyisi LSFien järjestettyyn
 ryhmään. Tämä on tavallisesti hyvin lievä pakotus määrättynä
 haulle. Näiden kahdenkymmenen seitsemän yhdistelmän opti-
 miyhdistelmä on valittu toimintalohkossa 71 perustuen sepst-
 raaliin särömittaan. Lopulta optimaalinen kategoria tai
 luokitus on määritetty myös sepstraalin särömitan perusteel-
 la. Kvantisoidut LSFT on muunnettu suodinkertoimiksi ja
 sitten autokorrelaatioviivästytykseksi interpolaatiotarkoituk-
 siin.

Lopputuloksena saatava LSF-vektroikvantisointikavio ei ole
 vain tehokas kaiuttimien kautta vaan myös erilaisten IRS--
 suodatusasteiden kautta, jotka mallintavat käsiyksikkösiir-
 timen vaikutusta. Vektroikvantisoijien koodikirjat on har-
 jaannutettu kuudenkymmenen puhujan puhetietokannasta käyt-
 täen tasaista sekä IRS-taajuusmuodostusta. Tämä on suunni-
 teltu antamaan yhdenmukaisen ja hyvän suorituskyvyn useiden
 kaiuttimien kautta ja erilaisten käsiyksiköiden kautta.

Keskimääräinen logaritminen spektrisäro koko TIA-puolen nopeuden tietokannan yli on noin 1,2 dB IRS-suodatetulle puhetiedolle ja noin 1,3 dB ei-IRS-suodatetulle puhetiedolle.

5

Kaksi viritysestimaattia on määritetty kahdesta viritysanalyysi-ikkunasta, jotka, kuten lineaariset ennustusanalyysi-ikkunat, on erotettu toisistaan 20 ms:llä. Ensimmäinen viritysanalyysi-ikkuna on keskitetty 40 ms:n kehyksen lopussa. Kukin viritysanalyysi-ikkuna on 301 näyttettä tai 37,625 ms pitkä. Viritysanalyysi-ikkunan poikkeama on esitetty KUV. 5B.

10

15

20

15
20
25
30
35

Viritysestimaatit lohkoissa 33 KUV. 3 on saatu viritysanalyysi-ikkunoista käyttämällä tunnetun viritysestimointialgoritmin modifioitua muotoa. Tunnetun virityksen seurantaalgoritmin vuokaavio on esitetty KUV. 7. Tämä virityksen estimointialgoritmi tekee alustavan viritysestimaatin toimintalohkoissa 73 käyttämällä peilifunktiota, joka on laskettu kaikille arvoille yksikössä {22.0, 22.5, ..., 114.5}. Tätä seuraa virityksen seuranta tuottamaan yleisen optimiviritysarvon. Taaksepäin katsova virityksen seuranta toimintalohkoissa 74 on tehty käyttämällä kahden aikaisemman viritysanalyysi-ikkunan virhefunktioita ja viritysestimaatteja. Eteenpäin katsova virityksen seuranta toimintalohkoissa 75 on tehty käyttämällä kahden tulevaisuuden viritysanalyysi-ikkunan virhefunktioita. Taksin ja eteenpäin katsovan virityksen seurantaan perustuvia viritysestimaatteja on verrattu päätöslohkoissa 76 tuottamaan yleisen optimiviritysarvon ulostulossa 77. Tunnettu virityksen estimointialgoritmi vaatii kahden tulevaisuuden viritysanalyysi-ikkunan virhefunktioita sen eteenpäin katsovalle virityksen seurannalle ja siten esittää 40 ms:n viiveen. Tämän epäkohdan välttämiseksi virityksen estimointialgoritmi on modifioitu keksinnöllä.

30

35

KUV. 8 esittää KUV. 3 avoimen silmukan virityksen estimaation 33 spesifisen toteutuksen. Virityksen analyysin puheikku-

na yksi ja kaksi ovat sissäntuloina vastaaville tietokonevirhefunktioille 331 ja 332. Näiden virhefunktioilaskelmien ulostulot ovat sissäntuloina viimeisten viritysestimaattien 333 puhdistamiselle, ja puhdistetut viritysestimaattit on lähetetty sekä taaksepäin että eteenpäin katsovaan virityksen seurantaan 334 ja 335 viritysikkuna yhdelle. Virityksen seurantapiirien ulostulot ovat sisääntuloina valitisimelle 336, joka valitsee avoimen silmukan virityksen yksi ensimmäiosenä ulostulona. Valittu avoimen silmukan viritys yksi on myös sisääntulona taaksepäin katsovalle virityksen seurantapiirille viritysikkuna kahdelle, joka antaa ulostulona avoimen silmukan virityksen kaksi.

KUV. 8 viritysestimointipiirillä toteutetun modifioidun virityksen seuranta-algoritmi on esitetty KUV. 9 vuokaaviossa. Modifioitu virityksen seuranta-algoritmi käyttää samaa virhefunktioita kuin tunnetussa virityksen estimointialgoritmissa kussakin viritysanalyysi-ikkunassa, mutta virityksen seurantakaavaa on muutettu.

20

Ennen virityksen seuranta joko ensimmäiselle tai toiselle viritysanalyysi-ikkunalle, kahden aikaisemman viritysanalyysi-ikkunan aikaisemmat kaksi viritysestimaattia on puhdistettu toimintalohkoissa 81 ja 82 vastaavasti, sekä taaksepäin että eteenpäin katsovan virityksen seurannan kanssa käyttämällä nykyisen kahden viritysanalyysi-ikkunan virhefunktioita. Tätä seuraa taaksepäin katsova virityksen seuranta toimintalohkossa 83 ensimmäiselle viritysanalyysi-ikkunalle käyttämällä kahden aikaisemman viritysanalyysi-ikkunan puhdistettuja viritysestimaatteja ja virhefunktioita. Eteenpäin katsova virityksen seuranta ensimmäiselle viritysanalyysi-ikkunalle toimintalohkossa 84 on rajoitettu käyttämällä toisen viritysanalyysi-ikkunan virhefunktioita. Kahta estimaattia verrataan päätöslohkossa 85 tuottamaan yleisesti parhaan viritysestimaatin ensimmäioselle viritysanalyysi-ikkunalle. Toiselle viritysanalyysi-ikkunalle taaksepäin katsova virityksen seuranta on suoritettu toimin-

talohkossa 86 sekä ensimmäisen viritysanalyysi-ikkunan viritysestimaatti ja sen virhefunktio. Eteenpäin katsovaa virityksen seurantaa ei ole käyetty tälle toiselle viritysanalyysi-ikkunalle tuloksena, että taaksepäin katsova viritysestimaatti on otettu olemaan yleisesti paras viritysestimaatti ulostulossa 87.

Joka 40 ms:n puhekehys on luokiteltu kahteen moodiin KUV. 3 lohossa 34. Yksi moodi on etupäässä ääneistetty ja on tunnettu hitaasti vaihtelevasta äänen alan muodosta ja hitaasti vaihtelevasta äänellisen soinnun värähtelynopeudesta tai virityksestä. Tätä moodia pidetään moodi A:na. Toinen moodi on pääasiassa ääneistämätön ja on omistettu moodi B:ksi. Moodin valinta perustuu sisääntuloihin, jotka on listattu alla:

1. Suodinkertoimet ensimmäiselle lineaariselle ennustusanalyysi-ikkunalle. Suodinkertoimiin viitataan $\{a(i)\}$:llä $0 \leq i \leq 10$ jossa $a(0)=1.0$,. Vektorimuodossa tähän on viitattu a_1 :nä.

2. Suodinkertoimien interpoloitu sarja ensimmäiselle lineaarisen ennustuksen analyysi-ikkunalle. Tämä interpoloitu sarja on saatu interpoloimalla kvantisoidut suodinkertoimet toiselle lineaarisen ennustuksen analyysi-ikkunalle nykyiselle 40 ms:n kehykselle ja aikaisemman 40 ms:n kehyksen vastaavalle autokorrelaatioalueella. Näihin suodinkertoimiin viitataan $\{\bar{a}_1(i)\}$:lla $0 \leq i \leq 10$ jossa $\bar{a}_1(0)=1.0$. Vektorimuodossa tähän on viitattu $\bar{a}_1$:na.

3. Aikaisemman toisen viritysanalyysi-ikkunan puhdistettu viritysestimaattiin on viitattu $\hat{P}_1$:llä.

4. Viritysestimaattiin ensimmäiselle viritysanalyysi-ikkunalle on viitattu P_1 :llä.

5. Viritysestimaatille toiselle viritysanalyysi-ikkunalle on viitattu P_2 :lla.

Käyttämällä ensimmäisiä kahta sisääntuloa sepstraalinen särön mitta $d_c(a_1, \bar{a}_1)$ suodinkertoimien $\{a_1(i)\}$ ja interpoloitujen suodinkertoimien $\{\bar{a}_1(i)\}$ välillä on laskettu ja ilmaistu dB:inä (desibeleinä). KUV. 3 moodivalitsimen 34 lohkokaavio on esitetty KUV. 10. Kvantisoidut suodinkertoimet lineaariselle ennustusikkunalle kaksi ja aikaisemman kehyksen lineaariselle ennustavalle ikkunalle kaksi ovat sisääntuloina interpolaattorille 341, joka interpoloi kertoimet autokorrelaatioalueella. Suodinkertoimien interpoloitu sarja on sisääntulona ensimmäisille kolmelle testipiirille. Tämä testipiiri 342 tekee suodinkertoimien interpoloidun sarjan testiin perustuvan sepstraalin särön ikkunalle kaksi vasten suodinkertoimia ikkunalle yksi. Toinen testipiiri 343 tekee aikaisemman viritysikkunan kaksi puhdistetun viritysestimaatin virityspoikkeamatestin vasten viritysikkunan yksi viritysestimaattia. Kolmas testipiiri 344 tekee viritysikkunan kaksi viritysestimaatin virityspoikkeaman testin vasten viritysikkunan yksi viritysestimaattia. Näiden testipiirien ulostulot ovat sisääntuloina moodivalitsimelle 345, joka valitsee moodin.

Kuten esitetty KUV. 11 vuokaaviossa, moodin valinta toteutettuna KUKV. 10 moodin määrittämisspiirillä on kolmivaiheinen prosessi. Ensimmäinen vaihe päätöslohkossa 91 on tehty sepstraalisen särömitan perusteella, jota verrataan annettuun absoluuttiseen kynnykseen. Jos kynnys on ylitetty, moodi on julistettu moodi B:ksi. Siten

$$STEP1: IF(d_c(a_1, \bar{a}_1) > d_{thresh}) Mode = ModeB$$

Tässä d_{thresh} on kynnys, joka on aikaisemman 40 ms:n kehyksen moodin funktio. Jos aikaisempi moodi oli moodi A, d_{thresh} saa arvon -6,25 dB. Jos aikaisempi moodi oli moodi B, d_{thresh} saa

5

$$STEP2: IF((1-f_{thresh}) \hat{P}_{-1} \leq P_1 \leq (1+f_{thresh}) \hat{P}_{-1}) Mode = ModeA.$$

10

STEP3: IF $((1 - f_{tresh}) P_2 \leq P_1 \leq (1 + f_{tresh}) P_2)$ Mode = ModeA.

20

Samaa kynnystekijää f_{tresh} käytetään molemmissa vaiheissa 2 ja 3. Lopulta, jos testi vaiheessa 3 epäonnistui, moodi nimitetään moodiksi B. Moodin valintaprosessin lopussa kynnykset d_{tresh} ja f_{tresh} päivitetään.

∴ 25

Moodille A toinen viritysestimaatti on kvantisoitu ja lähetetty, koska sitä on käytetty ohjaamaan suljetun silmukan viritysestimaattia kussakin alikehyksessä. Viritysestimaatin kvantisointi on suoritettu käyttämällä yhtenäistä 4-bitin kvantisoijaa. 40 ms:n puhekehys on jaettu seitsemään alikehykseen, kuten esitetty KUV. 12. Ensimmäiset kuusi ovat pituudeltaan 5,75 ms, ja seitsemäs on pituudeltaan 5,5 ms. Kussakin alikehyksessä syöttömalliparametrit on saatu sulje-

tun silmukan tavalla käyttämällä synteositekniikan analyysiä. Nämä KUV. 3 lohkoissa 35 käytetyt syöttömallin parametrit ovat adaptiivinen koodikirjaindeksi, adaptiivinen koodikirjavahvistus, kiinteä koodikirjaindeksi, kiinteä koodikirjavahvistus, ja kiinteä koodikirjavahvistusmerkki, kuten esitetty yksityiskohtaisemmin KUV. 13. Suodinkertoimet interpoloidaan autokorrelaatioalueella interpolaattorilla 3501, ja interpoloitu ulostulo on syötetty neljään kiinteään koodikirjaan 3502, 3503, 3504, ja 3505. Muut sisääntulot kiinteisiin koodikirjoihin 3502 ja 3503 syötetään adaptiivisella koodikirjalla 3506, kun taas muut sisääntulot kiinteisiin koodikirjoihin 3504 ja 3505 syötetään adaptiivisella koodikirjalla 3507. Kukin adaptiivisista koodikirjoista 3506 ja 3507 vastaanottaa sisääntulopuheen alikehykselle ja, vastaavasti, parametrit parhaalle ja toiseksi parhaalle polulle aikaisemmista alikehyksistä. Kiinteiden koodikirjojen 3502 - 3505 ulostulot ovat sisääntuloina vastaaville puhesynteesipiireille 3508 - 3511, jotka myös vastaanottavat interpoloidun ulostulon interpolaattorilta 3501. Piirien 3508 - 3511 ulostulot syötetään valitsimeen 3512, joka, käyttämällä signaali/kohinasuhteiden (SNR_t) mittaa, karsii ja valitsee parhaat kaksi polkua perustuen sisääntulopuheeseen.

Kuten esitetty KUV. 13, analyysi synteositekniikalla, jota käytetään saamaan syöttömalliparametrit, käyttää lyhyen ajan ennustajakertoimien interpoloitua sarjaa kussakin alikehyksessä. Syöttömalliparametrien optimaalisen sarjan määrääminen kullekin alikehykselle määrätään vain kunkin 40 ms:n kehyksen lopussa johtuen viivästetystä päätöksestä. Saataessa syöttömalliparametrit kaikkien seitsemän alikehyksen oletetaan olevan pituudeltaan 5,75 ms tai neljäkymmentäkuusi näytettä. Kuitenkin, viimeiselle tai seitsemännelle alikehykselle alikehyksen loppu päivittyy siten kuin adaptiivinen koodikirja päivittää ja paikallisten lyhytaikaisten ennustajatilamuuttujien päivitys suoritetaan vain alikehyksen pituudelle 5,5 ms tai neljäkymmentäneljä näytettä.

Lyhyen ajan ennustajaparametrit tai lineaariset ennustus-suodinparametrit interpoloidaan alikehyksestä alikehykseen. Interpolaatio suoritetaan autokorrelaatioalueella. Toiselle lineaariselle ennustusanalyysi-ikkunalle tarkoitettuista kvantisoiduista suodinkertoimista saatuihin normalisoituji-

5 hin autokorrelaatiokertoimiin viitataan $\{\rho_{-1}(i)\}$:llä aikaisemmalle 40 ms:n kehykselle ja $\{\rho_2(i)\}$:llä $(p_2(i))$:llä nykyiselle 40 ms:n kehykselle $0 \leq i \leq 10$ jossa $\rho_{-1}(0) = \rho_2(0) = 1,0$. Sitten interpoloidut autokorrelaatiokertoimet $\{\rho'_m(i)\}$ on sitten

10 annettu

$$\rho'_m(i) = v_m \cdot \rho_2(i) + [1 - v_m] \cdot \rho_{-1}(i), 1 \leq m \leq 7, 0 \leq i \leq 10,$$

tai vektorimuodossa

$$\rho'_m = v_m \cdot \rho_2 + [1 - v_m] \cdot \rho_{-1}, 1 \leq m \leq 7,$$

Tässä, v_m on interpolointipaino alikehykselle m . Interpoloidut viiveet $\{\rho'_m(i)\}$ on peräkkäin muunneltu lyhyen ajan ennustajasuodinkertoimiin $\{a'_m(i)\}$.

15

Interpolaatiopainojen valinta vaikuttaa äänen laatuun tässä moodissa huomattavasti. Tästä syystä ne täytyy määrittää huolellisesti. Nämä interpoloivat painot v_m on määritetty alikehykselle m minimoimalla keskimääräisen neliön virheen todellisen lyhyen ajan spektriverhokäyrän $S_{m,j}(\omega)$ ja interpoloidun lyhyen ajan tehon spektriverhokäyrän $S'_{m,j}(\omega)$ välillä kaikkien hyvin suuren puhetietokannan puhekehysien J yli. Toisin sanoen, E_m on määritetty minimoimalla

20

25

$$E_m = \sum_J \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_{m,j}(\omega) - S'_{m,j}(\omega)|^2 d\omega$$

Jos todellisiin autokorrelaatiokertoimiin alikehykselle m kehyksessä J on viitattu $\{\rho'_{m,j}(k)\}$, sitten määritelmällä

$$S_{m,j}(W) = \sum_{k=-10}^{10} P_{mj}(k) e^{-jwk}$$

$$S'_{mj}(\omega) = \sum_{k=-10}^{10} \rho'_{mj}(k) e^{-jwk}$$

Korvaamalla yllä olevat yhtälöt edelliseen yhtälöön, voidaan osoittaa, että minimoimalla E_m on ekvivalentti sille, että minimoidaan E'_m , jossa E'_m on annettu

$$E'_m = \sum_j \sum_{k=-10}^{10} [\rho_{mj}(k) - \rho'_{mj}(k)]^2$$

5

tai vektorimuodossa

$$E'_m = \sum_j \|\rho_{m,j} - \rho'_{m,j}\|^2,$$

jossa $\|\cdot\|$ edustaa vektroinormia. Korvaamalla ρ'_m yllä olevaan yhtälöön, erotamalla ν_m :n suhteen ja asettamalla se nolllaksi johtaa

$$\nu_m = \frac{[\sum_j \langle x_j, y_{m,j} \rangle]}{\sum_j \|x_j\|^2}$$

jossa $X_j = \rho_{2,j} - \rho_{-1,j}$ ja $Y_{m,j} = \rho_{m,j} - \rho_{-1,j}$ ja $\langle X_j, Y_{m,j} \rangle$ on pistetulo vektorien X_j ja $Y_{m,j}$ välillä. ν_m :n arvot laskettuina yllä esitetyllä menetelmällä käyttäen hyvin suurta puhetietokantaa on lisäksi hienoviritetty huolellisilla kuuntelutesteillä.

Kohdevektori t_{ac} adaptiiviselle koodikirjahaulle on suhteessa puhevektoriin s kussakin alikehyksessä $s = H t_{ac} + z$:lla. Tässä, H on neliötä alempi kolmiopmainen toeplitz-matriisi, jonka

ensimmäinen sarake sisältää interpoloidun lyhyen ajan ennustajan $\{a'_m(i)\}$ impulssivasteen alikehykselle m ja z on vektori sisältäen sen nollasisääntulovasteen. Kohdevektori t_{ac} on helpoimmin laskettu vähentämällä nollaimpulssivaste z puhevektorista s ja suodattamalla erotus käänteisellä lyhyen ajan ennustajalla, jossa on nolla-alkutilat.

Adaptiivinen koodikirjahaku adaptiivisissa koodikirjoissa 3506 ja 3507 käyttää spektrisesti painotettua keskimääräistä neliövirhettä ϵ_i mittaamaan erotuksen ehdokasvektorin r_i ja kohdevektorin t_{ac} välillä, kuten annettu

$$\epsilon_i = (t_{ac} - \mu_i r_i)^T W (t_{ac} - \mu_i r_i).$$

Tässä μ_i on liittyvä vahvistus ja W on spektripainotusmatriisi. W on positiivisesti määrätty symmetrinen toeplitz--matriisi, joka on saatu painotetun lyhyen ajan ennustajan typistetyistä impulssivasteista, jossa on suodinkertoimet $\{a'_m(i)\}$. Painotuskerroin γ on 0,8. Korvaamalla optimilla μ_i yllä olevassa yhtälössä, särötermi voidaan kirjoittaa

$$\epsilon_i = t_{ac}^T W t_{ac} - \frac{[\rho_i]^2}{e_i},$$

jossa ρ_i on korrelaatiotermi $t_{ac}^T W r_i$ ja e_i on energiatermi $r_i^T W r_i$. Vain ne ehdokkaat on huomioitu, joilla on positiivinen korrelaatio. Parhaat ehdokasvektorit ovat ne, joilla on positiiviset korrelaatiot ja suurimmat arvot

$$\frac{[\rho_i]^2}{e_i}.$$

Ehdokasvektori r_i vastaa erilaisiin viritysviiveisiin. Viritysviiveet näytteissä muodostuvat neljästä alialueesta. Ne ovat $\{20.00\}, \{20.5, 20.75, 21.0, 21.25, \dots, 50.25\}, \{50.50, 51.0, 51.5, 52.0, 52.5, \dots, 87.5\}$, ja $\{88.0, 89.0, 90.0,$

91.0, ..., 146.0}. On yhteensä 225 viritysviivettä ja vastaavaa ehdokasvektoria. Ehdokasvektori, joka vastaa kokonaislukuviivettä L , luetaan yksinkertaisesti soveltuvasta koodikirjasta, joka on kokoelma menneistä syöttönäytteistä.

5 Sekoitetulle (kokonaisluku plus murtoluku) viiveelle $L + f$, osa soveltuvasta koodikirjasta, joka on keskitetty alueelle, joka vastaa kokonaislukuviivettä L suodatetaan monivaiheisella suodattimella, joka vastaa murtolukua f . Epätäydelliset ehdokasvektorit, jotka vastaavat alhaisiin viiveisiin,

10 jotka ovat lähellä alialuetta tai sitä pienempiä, on tehty samalla tavalla kuin J. Campbell ym. ovat ehdottaneet. Monivaiheisen suodattimen kertoimet on saatu Hamming-ikkunoidusta sinkkifunktiosta. Kussakin monivaiheisessa suodattimessa on kuusitoista ulosotta.

15 Soveltuvan koodikirjan haku ei hae kaikkia ehdokasvektoreita. 6-bittinen hakualue on määritetty nykyisen 40 ms kehyksen kvantisoidun avoimen silmukan viritysestimaatilla P'_2 ja edellisen 40 ms kehyksen P'_1 , jos se oli moodin A kehys, vastaavalla. Jos edellinen moodi oli moodi B, silloin P'_1 on otettu olemaan viimeisen alikehyksen viritysviive edellisessä kehyksessä. Tämä 6-bittinen alue on keskitetty P'_1 :n lähelle ensimmäiselle alikehykselle ja P'_2 :n lähelle seitsemännelle alikehykselle. Väli-alikehyksille kahdesta kuuteen

20 6-bittinen halualue käsittää kaksi 5-bittistä halualuetta. Yksi on keskitetty P'_1 :n ja toinen P'_2 :n lähelle. Jos nämä kaksi aluetta menevät päällekkäin ja eivät ole toisensa poissulkevia, silloin käytetään hyväksi yhtä 6-bittistä aluetta, joka on keskitetty $(P'_1 + P'_2)/2$:n lähelle. Ehdokasvektori, jossa viritysviive on tällä alueella, on käännetty 6-bittiseksi indeksiksi. Nollaindeksi on varattu kaikki nollatulle soveltuvalle koodikirjavektorille. Tämä indeksi on valittu, jos kaikilla ehdokasvektoreilla hakualueella ei ole positiivisia korrelaatioita. Tämä indeksi on sovitettu virittämällä 6-bittinen tai kuudenkymmenen neljän viiveen

25 hakualue kuudenkymmenen kolmen viiveen hakualueeksi. Soveltu-

30

35

van koodikirjan vahvistus, joka on pakotettu olemaan positiivinen, on määriteltä hakusilmukan ulkopuolella ja on kvantisoitu käyttämällä 3-bittistä kvantisointitaulukkoa.

- 5 Koska käytetään viivästettyä päätöstä, soveltuva haku tuottaa kaksi parasta viritysviive- tai viritysjättämäehdokasta kaikissa alialueissa. Lisäksi, alikehyksille kahdesta kuuteen tämä on toistettava kahdelle parhaalle kohdevektorille, jotka tuotetaan kahdella parhaalla syöttömaööiparametrien
- 10 sarjalla, jotka on saatu edellisille alikehyksille nykyisessä kehyksessä. Tämä johtaa kahteen parhaaseen jättämäehdokkaaseen ja niihin liittyviin kahteen soveltuvaan koodikirjavahvistukseen alikehykselle yksi ja neljään parhaaseen jättämäehdokkaaseen ja niihin liittyviin neljään soveltuvaan
- 15 koodikirjavahvistukseen alikehyksille kahdesta kuuteen kunkin hakuprosessin lopussa. Kussakin tapauksessa kohdevektori kiinteälle koodikirjalle on saatu vähentämällä skaalattu soveltuva koodikirjavektori kohteesta soveltuvalle koodikirjahaulle, ts. $t_{ac} = t_{ac} - \mu_{opt} r_{opt}$, jossa r_{opt} on valittu soveltuva
- 20 koodikirjavektori ja μ_{opt} on siihen liittyvä soveltuva koodikirjavahvistus.

Moodissa A käytetään 6-bittistä glottaalisen pulssin koodikirjaa kiinteänä koodikirjana. Glottaalisen pulssin koodikirjavektorit on generoitu aika-siirrettyinä perusglottaalipulssin sekvensseinä, jotka on tunnettu parametreista, kuten paikka, vinous ja kesto. Glottaalinen pulssi on ensin laskettu 16 kHz:n näytteenottotaajuudella, kuten

$$g(n) = 0, 0 \leq n \leq n_0,$$

$$g(n) = A \cdot \sin^2 \left(\frac{\pi (n - n_0) T}{2T_p} \right), n_0 < n \leq n_0 + n_1,$$

$$g(n) = A \cdot \cos \left(\frac{\pi ((n - n_0) T - T_p)}{2T_n} \right), n_0 + n_1 < n \leq n_0 + n_2,$$

$$g(n) = 0, n_0 + n_2 < n \leq n_g,$$

Yllä mainituissa yhtälöissä eri parametrien arvot on oletettu olemaan $T = 62.5 \mu s$, $T_p = 440 \mu s$, $T_n = 1760 \mu s$, $n_0 = 88$, $n_1 = 7$, $n_2 = 35$, ja $n_g = 232$. Glottaalinen pulssi, joka on
 5 määritelty yllä, on differentioitu kahdesti tasoittamaan sen spektrimuotoa. Se on siten alipäästösuodatettu kolmekymmenellä kahdella ulosotolla varustetulla lineaarivaiheFIR-suodattimella, viritettynä 216 näytteen pituudelle, ja lopulta pienennetty 8 kHz:n näytteenottonopeudelle tuottamaan glottaalisen pulssin koodikirjan. Glottaalisen pulssin koodikirjan lopullinen pituus on 108 näytettä. Parametri A on sovitettu niin, että glottaalisen pulssin koodikirjan sisääntuloilla on tehollinen (RMS) arvo 0.5:a sisääntuloa kohti. Lopullinen glottaalisen pulssin muoto on esitetty KUV. 14.
 10 Koodikirjalla on 67.6 % niukkuus ensimmäisellä kolmellakymmenellä kuudella sisääntulolla ja viimeisellä kolmellakymmenelläseitsemällä sisääntulolla nolla.

On kuusikymmentäkolme glottaalista koodikirjavektoria, joista kukin on pituudeltaan neljäkymmentäkuusi näytettä. Kukin vektori on kartoitettu 6-bittiseen indeksiin. Nollas indeksi on varattu kaikille nolla olevalle kinteälle koodikirjavektorille. Tähän indeksiin on viitattu, jos hakutulokset vektorissa, mikä lisää säröä sen sijaan että se vähentäisi sitä. Jäljelle jäävät kuusikymmentäkolme indeksiä on omistettu kullekin kuudestakymmenestäkolmesta glottaalisen pulssin koodikirjavektorista. Ensimmäinen vektori muodostuu
 20
 25

ensimmäisistä neljästäkymmenestäkuudesta sisääntulosta koodikirjassa, toinen vektori muodostuu neljästäkymmenestäkuudesta sisääntulosta alkaen toisesta sisääntulosta, ja niin edelleen. Siten, on päällekkäin menoa, siirrettynä yhdellä, 67.6 % harvalla kiinteällä koodikirjalla. Lisäksi, ei-nollat elementit ovat koodikirjan keskellä, kun taas nollat ovat sen hännillä. Nämä kiinteän koodikirjan attributit on käytetty hyväksi sen haussa. Kiinteän koodikirjan haku käyttää samaa särömittausta kuin soveltuvassa koodikirjahaussa, jolla mitataan etäisyys kohdevektorin t_{sc} ja jokaisen kiinteän koodikirjan vektorin ehdokkaan välillä c_i , $ts: \xi_i = (t_{sc} - \lambda_i c_i)^T W (t_{sc} - \lambda_i c_i)$, jossa W on sama spektrinen painotusmatriisi, jota on käytetty soveltuvassa koodikirjahaussa. Vahvistuksen suuruus $|\lambda|$ on kvantisoitu kiinteän koodikirjan hakusilmukassa. Parittomille alikehyksille vahvistuksen suuruus on kvantisoitu käyttämällä 4-bittistä kvantisointitaulukkoa. Parillisille alikehyksille kvantisointi on tehty käyttämällä 3-bittistä kvantisointialuetta, joka on keskitetty edellisen alikehyksen kvantisoidun suuruuden lähelle. Tämä differentiaalinen vahvistuksen suuruuden kvantisointi ei ole vain tehokas bittien ehdoilla vaan myös vähentää monimutkaisuutta, koska tämä on tehty haun sisällä. Haku-
menetelmän lopussa säröä valitun koodikirjavektorin ja sen vahvistuksen kanssa on verrattu $t_{sc}^T W t_{sc}$:hen, säröön kaikille nollan kiinteän koodikirjan vektoreille. Jos särö on suurempi, sitten nollaindeksi on omistettu kiinteälle koodikirjan indeksille ja kaikki nollavektorit on otettu olemaan valittu kiinteä koodikirjavektori.

Viiivästetystä päätöksestä johtuen on kaksi kohdevektoria t_{sc} kiinteän koodikirjan haulle ensimmäisessä alikehyksessä vastaten kahteen parhaaseen jättämäehdokkaaseen ja niiden vastaaviin vahvistuksiin, jotka suljetun silmukan soveltuva koodikirjahaku antaa. Alikehyksille kahdesta seitsemään on neljä kohdevektoria vastaten kahta parasta syöttömalliparametrien sarjaa, jotka on määrätty edellisille alikehyksille

toistaiseksi ja kahdelle parhaalle jättämäehdokkaalle ja niiden vahvistuksille, jotka soveltuva koodikirjahaku antaa nykyisessä alikehyksessä. Kiinteän koodikirjan haku on sitten suoritettu kaksi kertaa alikehyksessä yksi ja neljä
 5 kertaa alikehyksissä kahdesta kuuteen, Mutta monimutkaisuus ei lisäännä suhteellisella tavalla, koska kussakin alikehyksessä energiatermiit $c_i^T W c_i$ ovat samat. Vain korrelaatiotermiit $t_{i,c}^T W c_i$ ovat erilaisia kussakin alikehyksen yksi kahdessa haussa ja kussakin neljässä haussa kahdesta seitsemään.

10

Viivästetty päätöshaku auttaa pehmentämään äänen korkeuden ja vahvistuksen ääri viivoja CELP-kooderissa. Viivästetty päätös on aikaansaatu tässä keksinnössä sellaisella tavalla, että yleinen koodekkiviive ei kasva. Siten, kussakin alikehyksessä, suljetun silmukan äänen korkeuden haku tuottaa M parasta estimaattia. Kullekin näistä M parhaasta estimaatista ja N parhaasta aikaisemmasta alikehysparametrasta on saatu MN optimaalista äänen korkeuden vahvistuksen indeksiiä, kiinteää koodikirjan indeksiiä, kiinteää koodikirjan vahvistusindeksiä, ja kiinteää koodikirjan vahvistusmerkkiä.

15

20

Alikehyksen lopussa nämä MN ratkaisua on karsittu L parhaaseen käyttämällä kumulatiivista SNR:ää nykyiselle 40 ms kehykselle kriteerinä. Ensimmäiselle alikehykselle käytetään $M = 2$, $N = 1$ ja $L = 2$. Viimeiselle alikehykselle käytetään $M = 2$, $N = 2$ ja $L = 1$. Kaikille muille alikehyksille käytetään $M = 2$, $N = 2$ ja $L = 2$. Viivästetyn päätöksen lähestymistapa on erikoisen tehokas ääneistettyjen siirrosta ääneistämättömiin ja ääneistämättömien siirrosta ääneistettyihin alueisiin. Tämä viivästetyn päätöksen lähestymistapa johtaa N kertaa suljetun silmukan äänenkorkeuden haun monimutkaisuuteen, mutta paljon vähemmän kuin MN kertaa kiinteän koodikirjan haun monimutkaisuus kussakin alikehyksessä. Tämä on siksi, että vain korrelaatiotekijät tarvitsee laskea MN kertaa kiinteälle koodikirjalle kussakin alikehyksessä, mutta energiatermiit täytyy laskea vain kerran.

...

...

Optimaaliset parametrit kullekin alikehykselle on määriteltä vain 40 ms kehyksen lopussa käyttämällä juovavarmistusta. MN ratkaisun karsiminen L ratkaisuun on varastoitu kullekin alikehykselle mahdollistamaan juovavarmennus. Esimerkki
 5 siitä, kuinka juovavarmennus on tehty, on esitetty KUV. 15. Tumma, paksu viiva ilmaisee optimaalisen tien, joka on saatu juovavarmistuksella viimeisen alikehyksen jälökeen.

Moodille B kumpaakin linjaspektrisen taajuusvektorin kvan-
 10 tisointiindeksien sarjaa ei tarvitse lähettää. Mutta kumpaa-
 kaan kahdesta avoimen silmukan äänenkorkeuden estimaatista ei ole lähetetty, koska niitä ei käytetä ohjattaessa suljetun silmukan äänenkorkeuden estimointia moodissa B. Tähän sisältyvä suurempi monimutkaisuus sekä lyhyen termin ennus-
 15 tajarparametrien suurempi bittinopeus moodissa B on kompensoitu syöttömalliparametrien hitaammalla päivityksellä.

Moodille B 40 ms puhekehys on jaettu viiteen alikehykseen. Kukin alikehys on 8 ms tai kuudenkymmenen neljän näytteen
 20 pituinen. Syöttömalliparametrit kussakin alikehyksessä ovat adaptiivinen koodikirjaaindeksi, adaptiivinen koodikirjavahvistus, kiinteä koodikirjaaindeksi, ja kiinteä koodikirjavahvistus. Ei ole kinteää koodikirjavahvistuksen merkkiä, koska se on aina positiivinen. Näiden parametrien parhaat estimaatit on määriteltä käyttämällä analyysi-synteesi-menetelmää
 25 kussakin alikehyksessä. Yleisesti paras estimaatti on määriteltä 40 ms kehyksen lopussa käyttämällä moodi A:n kanssa samanlaista viivästetyn päätöksen lähestymistapaa.

Lyhyen termin ennustajarparametrit tai lineaarisen ennustuk-
 30 sen suodinparametrit on interpoloitu alikehyksestä alikehykseen autokorrelaatioviivästyksen alueella. Toisen lineaarisen ennustusanalyysin kvantisoiduista suodinkertoimista saatujen normalisoitujen autokorrelaatioviivästyksiin on
 35 viitattu $\{\rho'_1(i)\}$:llä aikaisemmalle 40 ms kehykselle. Nykyisen 40 ms kehyksen ensimmäisten ja toisten lineaaristen ennustusanalyysi-ikkunoiden vastaaviin viivästyksiin on

viitattu $\{\rho_1(i)\}$:llä ja $\{\rho_2(i)\}$:llä vastaavsti. Normalilsointi takaa, että $\rho_1(0)=\rho_1(0)=\rho_2(0)=1.0$. Interpoloidut autokorrelaatioviivästyukset $\{\rho'_1(i)\}$ on annettu

$$\rho'_m(i) = \alpha_m \cdot \rho_1 + \beta_m \cdot \rho_1(i) + [1 - \alpha_m - \beta] \cdot \rho_2 \quad 1 \leq m \leq 5, 0 \leq i \leq 10,$$

5 tai vektroesityksenä

$$\rho'_m = \alpha_m \cdot \rho_1 + \beta_m \cdot \rho_1 + [1 - \alpha_m - \beta] \cdot \rho_2 \quad 1 \leq m \leq 5,$$

Tässä α_m ja β_m ovat interpolointipainot alikehykselle m . Interpolointiviivästyukset $\{\rho'_1(i)\}$ ovat peräkkäin konvertoituneet lyhyen termin ennustussuodinsparametrit $\{\alpha'_1(i)\}$.

10

Interpolointipainojen valinta ei ole niin kriittinen tässä moodissa kuin se on moodissa A. Kuitenkin ne on määritetty käyttämällä samoja objektiivisia kriteereitä kuin moodissa A ja hienovirittämällä niitä varovaisella mutta epävirallisilla kuuntelutesteillä. α_m ja β_m arvot, jotka minimoivat objektiiviset kriteerit E_m voidaan esittää olevan

15

$$A = \sum_J \|\rho_{-1,J} \rho_{2,J}\|^2$$

$$\beta_m = \frac{X_m C - Y_m A}{C^2 - AB}$$

jossa

$$A = \sum_J \|\rho_{-1,J} \rho_{2,J}\|^2$$

$$B = \sum_J \|\rho_{-1,J} \rho_{2,J}\|^2$$

$$C = \sum_J \langle \rho_{-1,J} \rho_{2,J} \rho_{1,J} \rho_{2,J} \rangle$$

$$X_m = \sum_J \langle \rho_{-1,J} - \rho_{2,J} \rho_{m,J} - \rho_{2,J} \rangle$$

$$Y_m = \sum_J \langle \rho_{m,J} - \rho_{2,J} \rho_{1,J} \rho_{1,J} - \rho_{2,J} \rangle$$

Kuten ennen $\rho_{-1,J}$ viittaa kehyksen J-1 toisen lineaarisen ennustusanalyysi-ikkunan kvantisoiduista suodinkertoimista saatuun autokorrelatioviivästysvektoriin, $\rho_{1,J}$ viittaa kehyksen J ensimmäisen lineaarisen ennustusanalyysi-ikkunan kvantisoiduista suodinkertoimista saatuun autokorrelaatioviivästysvektoriin, $\rho_{2,J}$ viittaa kehyksen J toisen lineaarisen ennustusanalyysi-ikkunan kvantisoiduista suodinkertoimista saatuun autokorrelaatioviivästysvektoriin, ja $\rho_{m,J}$ viittaa kehyksen J alikehyksessä m olevista puhenäytteistä saatuun todelliseen autokorrelaatioviivästysvektoriin.

Kiinteä koodikirja on kaksi osaa käsittävä 9-bittinen mini-innovatiivinen koodikirja. Yksi on Hadamar-vektorisummaosa ja toinen on yksittäsipulssiosa. Tämä koodikirja käyttää hakumenetelmää, joka käyttää hyväksi näiden osien rakenteen ja takaa positiivisen vahvistuksen. Tämä erikoinen koodikirja ja siihen liittyvä hakumenetelmä on D. Linin julkaisu "Ultra-fast Celp Coding Using Deterministic Multicodebook Innovations", ICASSP 1992, I 317 - 320.

Yksi mini-innovatiivisen koodikirjan komponentti on deterministinen vektorisummakoodi, joka on konstruoitu Hadamard-matriisilla H_m . Vektrosummakoodin koodivektori siten kuin se on käytetty tässä keksinnössä on ilmaistu

$$U_1 = \sum_{m=1}^4 \theta_m v_m(n), 0 \leq i \leq 15,$$

jossa perusvektorit $v_m(n)$ on saatu Hadamard-Sylvester-matriisin riveistä ja $\theta_m = \pm 1$. Perusvektorit on valittu perustuen

Hadamard-matriisin sekvenssijakoon. Hadamard-vektorisummakoodikirjojen koodivektorit ovat arvoja ja binääriarvioituja koodisekvenssejä. Verrattuna aikaisemmin harkittuihin algebrallisiin koodeihin, Hadamard-vektorisummakoodit on konstruoitu omistamaan ideaalisemmat taajuus- ja vaiheominaisuudet. Tämä johtuu Hadamard-matriisille tässä keksinnössä käytetystä perusvektorin osituskaaviosta, joka voidaan tulkita yhdenmukaisena näytteenottona sekvenssijärjestetystä Hadamard-matriisin rivivektoreista. Vastakohtana, ei-yhdenmukaiset näytteenottomenetelmät ovat tuottaneet huonompia tuloksia.

Moni-innovatiivisen koodikirjan toinen komponentti ovat yksittäiset pulssikoodisekvenssit käsittäen aikasiirretyn deltaimpulssin sekä yleisemmät syöttöpulssimuodot konstruoituina diskreeteistä sinc- ja cosc-funktioista. Yleistetyt pulssimuodot on määritelty

$$z_1(n) = A \text{sinc}(n) + B \text{cosc}(n+1),$$

ja

$$z_1(n) = A \text{sinc}(n) + B \text{cosc}(n+1),$$

jossa

$$\text{sinc}(n) = \frac{\sin(\pi n)}{\pi n}, n \neq 0, \text{sinc}(0) = 1$$

ja

$$\text{cosc}(n) = \frac{1 - \cos(\pi n)}{\pi n}, n \neq 0, \text{cosc}(0) = 1$$

kun sinc- ja cosc-funktiot ovat aikapoikkeutettuja, ne vastaavat sitä, mikä tunnetaan sinc-perusfunktiona $z_0(n)$. Epäviralliset kuuntelutestit näyttävät, että aikasiirretyt pulssimuodot parantavat syntetisoidun puheen äänenlaatua.

Kiinteän koodikirjan vahvistus on kvantisoitu käyttämällä neljää bittiä kaikissa alikehyksissä hakusilmukan ulkopuolella. Kuten aikaisemmin korostettiin, vahvistuksen on taattu olevan positiivinen ja siksi ei merkkibittiä tarvitse lähettää kullakin kiinteällä koodikirjan vahvistusindeksillä. Johtuen viivästetystä päätöksestä alikehyksessä yksi on kaksi sarjaa optimaalisia kiinteän koodikirjan indeksejä ja vahvistuksia ja neljä sarjaa alikehyksissä kahdesta viiteen.

Viivästetyn päätöksen lähestymistapa moodissa B on identtinen siihen, jota käytetään moodissa A. Optimaaliset parametrit kullekin alikehykselle on määritetty 40 ms kehyksen lopussa käyttämällä identtistä jovavarmistuksen menetelmää.

Puhdedekooderi 46 (FIG. 4) on esitetty KUV. 16 ja vastaanottaa vaimennetun puheen bittivirran samassa muodossa kuin annettu ulos pudekooderilla KUV. 18. Parametrit ovat pakkaamattomina sen määrittämisen jälkeen, onko vastaanotetun moodin bitti (ensimmäisen vaimennetun sanan MSB) 0 (moodi A) tai 1 (moodi B). Näitä parametrejä käytetään sitten syntetisoimaan puhetta. Lisäksi puhdedekooderi vastaanottaa syklisen vapautustarkistuksen (CRC) perustuen huonon kehyksen ilmaisimeen kanavadekooderista 45 (KUV. 1). Tätä huonon kehyksen ilmaisimen lippua käytetään laukaisemaan dekooderin huonon kehyksen virheen maskintamisen ja virheen paljastamisen osat (ei esitetty). Nämä voidaan myös laukaista joillain sisäänrakennetuilla virheelmaisukaavioilla.

KUV. 9 moodille A linjaspektraalisen taajuusvektorin kvantisointi-indeksien toista sarjaa käytetään antamaan osoite kiinteälle koodikirjalle 101 kvantisoitujen suodinkertoimien rekonstruoimiseksi. Kiinteän koodikirjan vahvistuksen bitit sisäänsyötettyinä skaalaavaan kertojaan 102 muuttavat kvantisoidut suodinkertoimet autokorrelattioviivästeiksi interpolaatiotarkoituksiin. Kussakin alikehyksessä autokorrelattioviivästyksset on interpoloitu ja muunnettu lyhyen ajan ennustajakertoimiksi. Perustuen avoimen silmukan kvantisoi-

tuun äänenkorkeuden estimaattiin kertojasta 102 ja suljetun silmukan äänenkorkeuden indeksiin kertojasta 104, absoluuttinen äänenkorkeuden viivearvo on määriteltä kussakin alikehyksessä. Vastaava vektori adaptiivisesta koodikirjasta 103 on skaalattu sen vahvistuksella skaalaavassa kertojassa 104 ja summattu summaimella 105 skaalatun kiinteän koodikirjan vektorin kanssa tuottamaan syöttövektori kussakin alikehyksessä. Tätä syöttösignaalia käytetään suljetun silmukan ohjauksessa, ilmaistuna katkoviivalla 106, antamaan osoite adaptiiviseen koodikirjaan 103. SDyöttösignaali on myös äänenkorkeuden esisuodatettu suodattimessa 107, kuten I.A. Gerson ja M.A. Jasuik, *supra*, ovat kuvanneet ennen puhesynteesiä käyttämällä lyhyen termin ennustajaa, jossa on interpoloidut suodinkertoimet. Äänen korkeuden suodattimen 107 ulostulo on edelleen suodatettu synteesisuodattimessa 108, ja tuloksena oleva syntetisoitu puhe on korotettu kläyttämällä globaalia napanollajäölisuodatinta 109, jota seuraa spektraalinen kaltevuuden korjaava yksinapainen suodatin (ei esitetty). Jälkisuodatetun puheen energian normalisointi on loopullinen vaihe.

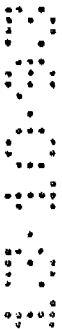
Moodille B, molempia juovaspektraalisen taajuuden vektorin kvantisointiindeksien sarjoja käytetään rekonstruoimaan autokorrelaatioviivästyksen sekä ensimmäiset että toiset sarjat. Kussakin alikehyksessä autokorrelaatioviivästyksset on interpoloitu ja muunnettu lyhyen termin ennustajakertoimiksi. Syöttövektori kussakin alikehyksessä on rekonstruoitu yksinkertaisesti kuten skaalattu adaptiivinen koodikirjavektori koodikirjasta 103 lisättynä skaalatulla kiinteällä koodikirjavektorilla koodikirjasta 101. Syöttösignaali on äänenkorkeuden suhteen esisuodatettu suodattimessa 107 kuten moodissa A ennen puhesynteesiä käyttämällä lyhyen termin ennustajaa, jossa on interpoloidut suodinkertoimet. Syntetisoitu puhe on myös korostettu käyttämällä samaa globaalia suodatinta 109, jota seuraa jälkisuodatetun puheen energian normalisointi.

Rajoitettu sisäänrakennettu virheenilmaisukyky on rakennettu dekodderiin. Lisäksi ulkoinen virheenilmaisuus on tehty sata-vaksi kanavadekodderista 45 (KUV. 4) huonon kehysilmaisimen lipun muodossa. Erilaisia virheentoipumiskaavioita käytetään erilaisille parametreille virheilmaisun tapauksessa. Moodibitti on selvästi herkin bitti ja tästä syystä se on sisällytetty huomiokyvyltään merkittävimpiin bitteihin, jotka vastaanottavat CRC-suojauksen ja varustettu puolen nopeuden suojauksella ja myös asennoilla seuraten konvolutionaalisen kooderin jälkibittejä maksimiimmunitetille. Lisäksi parametrit on pakattu vaimennettuun bittivirtaan siten, että, jos olisi virhe moodibitissä, sitten LSF VQ:n indeksien toinen sarja ja jotkut koodikirjan vahvistusindeksit voidaan yhä pelastaa. Jos moodibitti on virheessä, huonon kehysilmaisulippu asetettaisiin, mikä johtaa kaikkien virheen toipumismekanismien laukaisuun, joka johtaa asteettaiseen myksitämiseen. Sisäänrakennetut ilmaisukaaviot lyhyen termin ennustajaparametreille käyttävät hyväkseen tosiasiaa, että virheiden poissaollessa vastaanotetut LSFT on järjestetty.

Virheen toipumiskaaviot käyttävät interpolaatiota virheen tapauksessa vastaanotettujen LSFien ensimmäisessä sarjassa ja toistamista virheiden tapauksessa molempien LSFien sarjojen toisessa sarjassa. Kussakin alikehyksessä virheenlievenyskaavio virheen tapauksessa äänenkorkeuden viiveessä tai koodikirjan vahvistuksissa sisältää aikaisempin alikehysten arvojen toiston, mitä seuraa vahvistusten pieneneminen. Sisäänrakennettu ilmaisukyky on olemassa vain kiinteälle koodikirjan vahvistukselle ja se käyttää hyväksi sitä tosiasiaa, että sen suuruus harvoin heilahtaa yhdestä ääriarvosta toiseen alikehyksestä toiseen. Lopulta energiaperustaista virheenilmaisua juuri jälisuodattimen jälkeen käytetään tarkistuksena varmistamaan, että jälisuodatetun puheen energia kussakin alikehyksessä ei koskaan ylitä kiinteää kynnysarvoa.

Vaikka keksintöä on kuvattu yhden edullisen sovellutusmuodon puitteissa, alan ammattimiehet voivat havaita, että keksin-

töä voidaan toteuttaa soveltamalla oheisten vaatimusten hengessä ja laajuudessa.



PATENTTIVAATIMUKSET

1. Järjestelmä äänitiedon vaimentamiseksi, **tunnettu** siitä, että se käsittää:

5

laitteet (31) äänitiedon vastaanottamiseksi ja tiedon jakamiseksi äänikehyksille;

10

lineaarinen ennustava koodianalysointilaite ja kvantisoija (32), joka toimii tiedolla kussakin äänikehyksessä suorittamaan lineaarista ennustavaa koodianalyysiä ensimmäisessä ja toisessa ääni-ikkunassa, joka ensimmäinen ikkuna on keskitetty olennaisesti keskelle ja toinen ikkuna olennaisesti reunaan äänikehystä tuottamaan ensimmäiset ja toiset suodinkertoimien sarjat ja juovaspektritaajuusparit;

15

koodikirja sisältäen vektorikvantisointi-indeksin;

20

äänenkorkeuden estimaattorin (33) tuottamaan kaksi äänenkorkeuden estimaattia käyttämällä kolmatta ja neljättä äänenkorkeuden ikkunaa, jotka, kuten ensimmäinen ja toinen ikkuna, on vastaavasti keskitetty olennaisesti äänikehysten keskelle ja reunaan;

25

30

moodin määrääjän (34), joka vastaa ensimmäiseen ja toiseen suodinkertoimeen ja kahteen äänenkorkeuden estimaattiin äänikehysten luokitteluksi ensimmäiseen vallitsevasti ääneistettyyn moodiin; ja

35

lähetin (16) lähettämään ensimmäisen juovaspektritaajuuden vektorikvantisointikoodikirjan indeksien sarjan koodikirjasta ja toisen äänenkorkeuden estimaatin ohjaamaan suljetun silmukan äänenkorkeuden estimointia ensimmäisen moodin äänelle.

2. Patenttivaatimuksen 1 mukainen järjestelmä, **tunnettu** siitä, että se edelleen käsittää:

CELP-syöttöanalyysaattorin ohjaamaan suljetun silmukan äänenkorkeudenhakua ensimmäisessä moodissa;

5 viivästetyn päätöksen öaitteet puhdistamaan syöttömallin parametrir ensimmäisessä moodissa siten, että kokonaisviiveeseen ei vaikuteta; ja

10 kooderilaitteet (26) ensimmäiselle moodille jakamaan vastaanotetun äänikehyksen useaan alikehykseen ja kullekin alikehykselle määräämään äänenkorkeuden indeksin, äänenkorkeuden vahvistuksen indeksin, kiinteän koodikirjan indeksin, kiinteän koodikirjan vahvistuksen indeksin, ja kiinteän koodikirjan vahvistuksen merkin käyttämällä suljetun silmukan analyysi-synteesi-lähestymistapaa, jotka kooderilaitteet
15 suorittavat suljetun silmukan äänenkorkeuden indeksihaun keskitettynä olennaisesti kvantisoidun äänenkorkeuden estimaatin ympärille, joka on saatu toisesta nykyisen äänikehyksen sekä aikaisemman äänikehyksen äänenkorkeuden analyysi--ikkunasta.

20

3. Järjestelmä äänitiedon vaimentamiseksi, **tunnettu** siitä, että se käsittää:

25 laitteet (31) äänitiedon vastaanottamiseksi ja tiedon jakamiseksi äänikehyksille;

30 lineaarisen ennustavan koodianalyysaattorin ka -kvantisoijan (32), joka toimii tiedolla kussakin äänikehyksessä suorittamaan lineaarista ennustavaa koodianalyysiä ensimmäiselle ja toiselle ääni-ikkunalle, joka ensimmäinen ikkuna on keskitetty olennaisesti äänikehyksen keskelle ja toinen olennaisesti sen reunaan, tuottamaan ensimmäisen ja toisen sarjan suodinkertoimia ja juovaspektritaajuuspareja;

35 koodikirjan sisältäen vektorikvantisointi-indeksin;

5 äänenkorkeuden estimaattorin (33) tuottamaan kaksi äänenkorkeuden estimaattia käyttämällä kolmatta ja neljättä ääni-ikkunaa, jotka, kuten ensimmäinen ja toinen ikkuna, on vastavasti keskitetty olennaisesti äänikehyksen keskelle ja reunaan;

10 moodimäärittäjän (34), joka vastaa ensimmäistä ja toista suodinkerrointa ja kahta äänenkorkeuden estimaattia äänikehyksen luokitteluksi toiseen vallitsevasti ääneistettyyn moodiin; ja

lähettimen (16) lähettämään molemmat juovaspelktritaajuusvektorin kvantisointikoodikirjan indeksien sarjat.

15 4. Vaatimuksen 3 mukainen järjestelmä, **tunnettu** siitä, että se käsittää:

viivästetyn päätöksen laitteet jalostamaan syöttömalliparameterit siten, että kokonaisviiveeseen ei vaikuteta; ja

20

25 kooderilaitteet (26) jakamaan vastaanotetun äänikehyksen useaan alikehykseen ja määrittämään kullekin alikehykselle äänenkorkeuden indeksin, äänenkorkeuden vahvistusindeksin, kiinteän koodikirjan indeksin, ja kiinteän koodikirjan vahvistusindeksin käyttämällä suljetun silmukan analyysi-synteesi-lähestymistapaa, kooderilaitteet suorittamaan suljetun silmukan äänenkorkeuden indeksihaun joka on keskitetty olennaisesti kvantisoidun äänenkorkeuden estimaatin ympärille, joka on saatu toisesta nykyisen äänikehyksen äänenkorkeuden analyysistä sekä aikaisemman äänikehyksen viimeisen alikehyksen äänenkorkeuden analyysistä.

30 5. Jonkin vaatimuksen 1 tai 2 mukainen järjestelmä, **tunnettu** siitä, että se käsittää jonkin vaatimuksen 3 tai 4.

35 6. Vaatimuksen 1, 2 tai 5 mukainen järjestelmä, **tunnettu** siitä, että äänenkorkeuden estimaattori käsittää:

ensimmäiset ja toiset laskentalaitteet (331,332) vastaavsti vastaanottaen tietoa kolmannesta ja neljännestä ääni-ikkunasta virhefunktion laskemiseksi;

5 laitteet (331) vastaanottamaan lasketut virhefunktiot ensimmäisistä ja toisista laskentalaitteista jalostamaan menneet äänenkorkeuden estimaatit;

10 taakse- ja eteenpäin katsovat äänenkorkeuden jäljittäjät (337) vastaten jalostettuja menneitä äänenkorkeuden estimaatteja tuottamaan optimaaliset ensimmäiset ja toiset optimaaliset äänenkorkeuden estimaatit;

15 äänenkorkeuden valitsimet valitsemaan yhden ensimmäisestä ja toisesta optimaalisesta äänenkorkeuden estimaatista yhtenä äänenkorkeuden kahdesta estimaatista; ja

20 taaksepäin katsovan äänenkorkeuden jäljittäjän vastaten äänenkorkeuden valitsimeen antamaan ulos toisen äänenkorkeuden kahdesta estimaatista.

7. Vaatimuksen 6 mukainen järjestelmä, **tunnettu** siitä, että m,oodin määrittäslaitteet käsittävät:

25 ensimmäisen koeistajan (342) vastaanottamaan suodinkertoimien interpoloidun sarjan toiselle ikkunalle ja suodinkertoimet ensimmäiselle ikkunalle vertaamaan sepstraalista särömittaa kynnyksarvoon;

30 toisen koeistimen (343) vertaamaan jalostettua äänenkorkeuden estimaattia ikkunaan neljä ja äänenkorkeuden estimaattia ikkunalle kolme;

35 kolmannen koeistimen (344) vertaamaan äänenkorkeuden estimaattia ikkunalle neljä ja äänenkorkeuden estimaattia ikkunalle kolme; ja

moodiselektorin (345) valitsemaan ensimmäisen moodin, jos toisella ja kolmannella koestimella tehdyt vertailut ovat lähellä, mutta valitsemaan toisen moodin, jos ensimmäisellä koeistimella tehty vertailu ylittää kynnyksen, jonka arvo on aikaisemman moodin funktio.

8. Menetelmä vaimentamaan äänitietoa, tunnettu siitä, että se käsittää vaiheet:

10 äänitiedon (31) vastaanottaminen ja tiedon jakaminen äänikehyksiin;

15 lineaarisen ennustavan koodianalyysin (32) suorittamionen tiedolle ensimmäisessä ja toisessa ääni-ikkunassa kussakin äänikehyksessä; joka ensimmäisen ikkuna on keskitetty olenniasesti keskelle ja toinen ikkuna olennaisesti reunaan äänikehystä, ja suodinkertoimien ensimmäisen ja toisen sarjan tuottaminen;

20 äänenkorkeuden (33) kahden estimaatin tuottamionen käyttämällä kolmatta ja neljättä ääni-ikkunaa, jotka, kuten ensimmäinen ja toinen ikkuna, on vastaavasti keksitetty olennaisesti keskelle ja reunaan äänikehystä;

25 äänikehysten (34) luokitteleminen ensimmäiseen vallitsevasti ääneistettyyn moodiin, joka perustuu ensimmäiseen ja toiseen suodinkertoimeen ja kahteen äänenkorkeuden estimaattiin;

30 jovaspektritaajuusvektorin kvantisointi-indeksien toisen sarjan ja toisen äänenkorkeuden estimaatin lähettämisen ensimmäisessä moodissa ohjaamaan suljetun silmukan äänenkorkeuden estimointia.

35 9. Vaatimuksen 8 mukainen menetelmä, tunnettu siitä, että se edelleen käsittää:

syöttömalliparametrien jalostaminen ensimmäisessä moodissa käyttämällä viivästetyn päätöksen laitteita siten, että kokonaisviiveeseen ei vaikuteta;

- 5 äänikehyksen jakaminen ensimmäiseen alikehysten joukkoon ja määrittäminen kullekin alikehykselle äänenkorkeuden indeksi, äänenkorkeuden vahvistuksen indeksi, kiinteän koodikirjan indeksi, kiinteän koodikirjan vahvistuksen indeksi, ja kiinteän koodikirjan vahvistuksen merkki käyttämällä suljetun silmukan analyysi-synteesi-lähestymistapaa, kun kehys on
10 tunnistettu ensimmäisen moodin lähetyksenä; ja

- suljetun silmukan äänenkorkeuden indeksiaun (26) suorittaminen käyttämällä CELP-syöttöanalyysiä, joka on keskitetty
15 kvantisoidun äänenkorkeuden estimaatin ympärille, joka on saatu nykyisen äänikehyksen sekä aikaisemman äänikehyksen toisesta äänenkorkeuden analyysi-ikkunasta.

10. Mentelmä äänitiedon vaimentamiseksi, **tunnettu** siitä,
20 että se käsittää vaiheet:

äänitiedon (31) vastaanottaminen ja tiedon jakaminen äänikehyksiin;

- 25 lineaarisen ennustavan koodianalyysin (32) suorittaminen tiedolle ensimmäisessä ja toisessa ääni-ikkunassa kussakin äänikehyksessä, joka ensimmäinen ikkuna on keskitetty keskelle ja toinen ikkuna olenniasesti reunaan äänikehystä, ja suodinkertoimien ensimmäisen ja toisen sarjan tuottaminen;

- 30 äänenkorkeuden (33) kahden estimaatin tuottaminen käyttämällä kolmatta ja neljättä ääni-ikkunaa, jotka, kuten ensimmäinen ja toinen ikkuna, on vastaavasti keskitetty olennaisesti äänikehyksen keskelle ja reunaan;

35

äänikehyksen (311) luokitteleminen toiseen vallitsevasti ääneistettyyn moodiin perustuen ensimmäiseen ja toiseen suuodinkertoimeen ja äänenkorkeuden toiseen estimaattiin; ja

- 5 juovaspektritaajuusvektorin kvantisointi-indeksien molempien sarjojen lähettäminen toisessa moodissa.

11. Vaatimuksen 10 mukainen menetelmä, **tunnettu** siitä, että se edelleen käsittää:

10

syöttömalliparametrien jalostamisen toisessa moodissa käyttämällä viivästetyn päätöksen laitteita siten, että kokonaisviiveeseen ei vaikuteta;

- 15 äänikehyksen jakaminen ensimmäiseen alikehysten joukkoon ja määrittäminen kullekin alikehykselle äänenkorkeuden indeksi, äänenkorkeuden vahvistuksen indeksi, kiinteään koodikirjan indeksi, kiinteän koodikirjan vahvistusindeksi, käyttämällä suljetun silmukan analyysi-synteesi-lähestymistapaa, kun
20 kukin kehys on tunnistettu toisena moodilähetystenä; ja

25

suljetun silmukan äänenkorkeuden indeksihaun (26) suorittaminen keskitettynä kvantisoidun äänenkorkeuden estimaatin ympärille, joka on saatu nykyisen äänikehyksen toisesta äänenkorkeuden analyysistä sekä aikaisemman äänikehyksen viimeisen alikehyksen toisesta äänenkorkeuden analyysi-ikkunasta.

30

12. Vaatimuksen 8 tai 9 mukainen menetelmä, **tunnettu** siitä, että se sisältää jonkin vaatimuksista 10 tai 11.

35

Vaatimuksen 8, 9 tai 12 mukainen menetelmä, **tunnettu** siitä, että äänenkorkeuden estimaattien tuottaminen käsittää vaiheet:

40

tiedon vastaanottaminen kolmannesta ja neljännestä ääni-ikkunasta laskemaan virhefunktio;

laskettujen virhefunktioiden vastaanottaminen ensimmäisitä ja toisista laskentalaitteista jalostamaan menneet äänenkorkeuden estimaatit;

- 5 optimaalisten ensimmäisen ja toisen optimaalisen äänenkorkeuden estimaatin tuottaminen; ja

yhden valitseminen ensimmäisestä ja toisesta optimaalisesta äänenkorkeuden estimaatista yhtenä äänenkorkeuden kahdesta
10 estimaatista.

14. Vaatimuksen 8, 9 tai 12 mukainen menetelmä, tunnettu siitä, että äänikehyksen luokitteluvaihe sisältää vaiheet:

- 15 suodinkertoimien interpoloidun sarjan vastaanottaminen toiselle ikkunalle ja suodinkertoimien ensimmäiselle ikkunalle vertaamaan sepstraalista särömittaa kyynysarvoon;

20 jalostetun äänenkorkeuden vertaaminen ikkunaan neljä ja äänenkorkeuden estimaatin ikkunaan kolme;

äänenkorkeuden estimaatin vertaaminen ikkunaan neljä ja äänenkorkeuden estimaatin ikkunaan kolme; ja

- 25 ensimmäisen moodin valitseminen, jos äänenkorkeuden estimaatin vertaliuvaiheilla tehtyt vetailut ovat lähellä, mutta toisen moodin valitseminen, jos sepstraalin särön vertailuvaiheella tehty vertailu ylittää kynnyksen, jonka arvo on aikaisemman moodin funktio.

30 15. Vaatimuksen 9, 11 tai 12 mukainen menetelmä, tunnettu siitä, että suljetun silmukan äänenkorkeuden indeksihaun suorittamisvaihe käyttämällä CELP-syöttöanalyysiä käsittää seuraavat vaiheet:

- 35 äänikehyksen jakaminen joukkoon alikehyksiä ja kullekin alikehykselle suodinkertoimien interpoloidun sarjan käyttä-

minen käytettäväksi suljetun silmukan äänenkorkeuden haussa ja kiinteään koodikirjan haussa, joka interpolaatio on suoritettu viivästyksellä käyttäen interpolaatiopainojen optimaalista sarjaa;

5

kullekin alikehykselle syöttömalliparametrien määrittäminen vastaten syöttömalliparametrien kahta parasta sarjaa, jotka on määritetty aikaisemmissa alikehyksissä siihen asti; ja

10

edelleen käsittäen minkä tahansa tai mitkä tahansa seuraavista vaiheista:

15

kullekin alikehykselle ensimmäisessä moodissa tai kullekin syöttömalliparametrien sarjalle määritettynä tähän asti kahden optimaalisen suljetun silmukan äänenkorkeuden estimaatin määrittäminen hakemalla joukon äänenkorkeuden arvoja ei-tasaisesta äänenkorkeuden viivetaulukosta, joka on saatu kvantisoidusta avoimen silmukan äänenkorkeuden arvosta;

20

kullekin alikehykselle toisessa moodissa ja kullekin syöttömalliparametrien sarjalle määrättyä tähän asti, kahden optimaalisen suljetun silmukan äänenkorkeuden estimaatin määrittäminen hakemalla vain kokonaislukuisia äänenkorkeuden viiveitä;

25

kullekin alikehykselle ensimmäisessä moodissa ja kullekin syöttömalliparametrien sarjalle määrättyä tähän asti ja kullekin optimaaliselle suljetun silmukan äänenkorkeuden arvolle, glottaalisen pulssikoodikirjan hakeminen optimaaliselle glottaalipulssivektorille ja sen vahvistukselle, ja koodikirjan erikoisen rakenteen hyväksikäyttäminen haun aikana sekä sen tosiasian, että haussa käytetyt energiatermit tarvitsee laskea vain kerran;

30

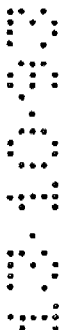
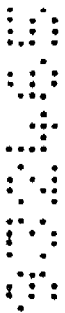
35

kullekin alikehykselle toisessa moodissa ja kullekin sarjalle syöttömalliparametrejä kääritettynä tähän asti, ja kullekin optimaaliselle suljetun silmukan äänenkorkeuden arvolle,

moni-innovatiivisen koodikirjan haku käyttämään optimaalista innovaatiosekvenssiä ja sen vahvistusta;

5 kunkin alikehyksen lopussa, lukuunottamatta ensimmäistä,
neljän syöttömalliparametrin sarjan karsiminen, johtaen
kahden aikaisemman sarjan ja kahden optimaalisen äänenkor-
keusestimaatin yhdistelmsätä kahteen syöttömalliparametrin
sarjaan käyttämällä kumulatiivista SNR:ää kriteerinä; ensim-
mäiseölle alikehykselle kahden syöttömalliparametrin sarjan
10 valitseminen vastaten vain parasta aikaisemmin määrätyn
syöttömalliparametrin sarjaa; tai

kunkin kehyksen lopussa, käyttämällä viivästetyn päätäksen
laitteita, optimaalisten syöttömalliparamteriindeksien
15 määrittäminen kussakin alikehyksessä tekemällä jäljitysperä.



13.10.83 8:36:33

FIG. 1

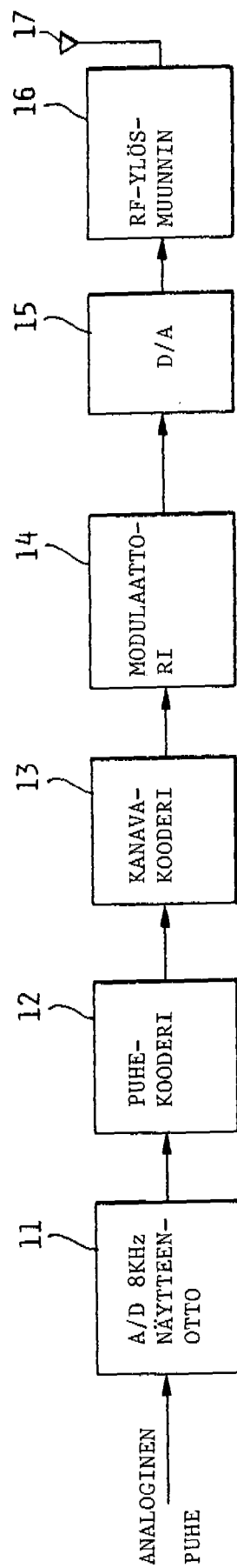
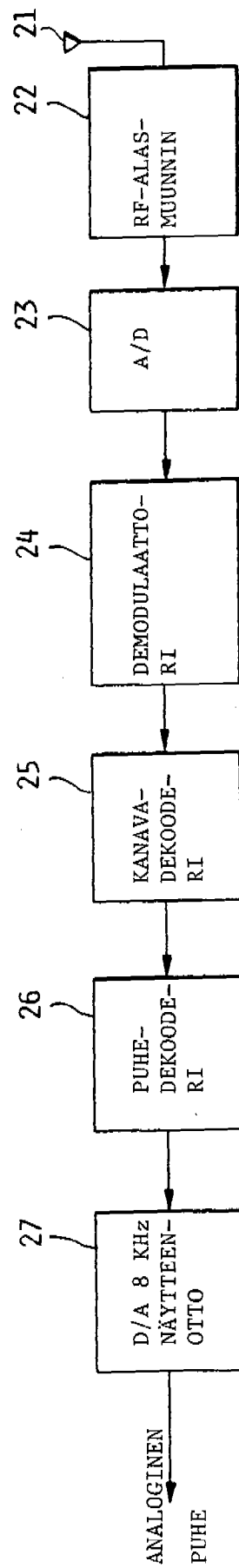


FIG. 2



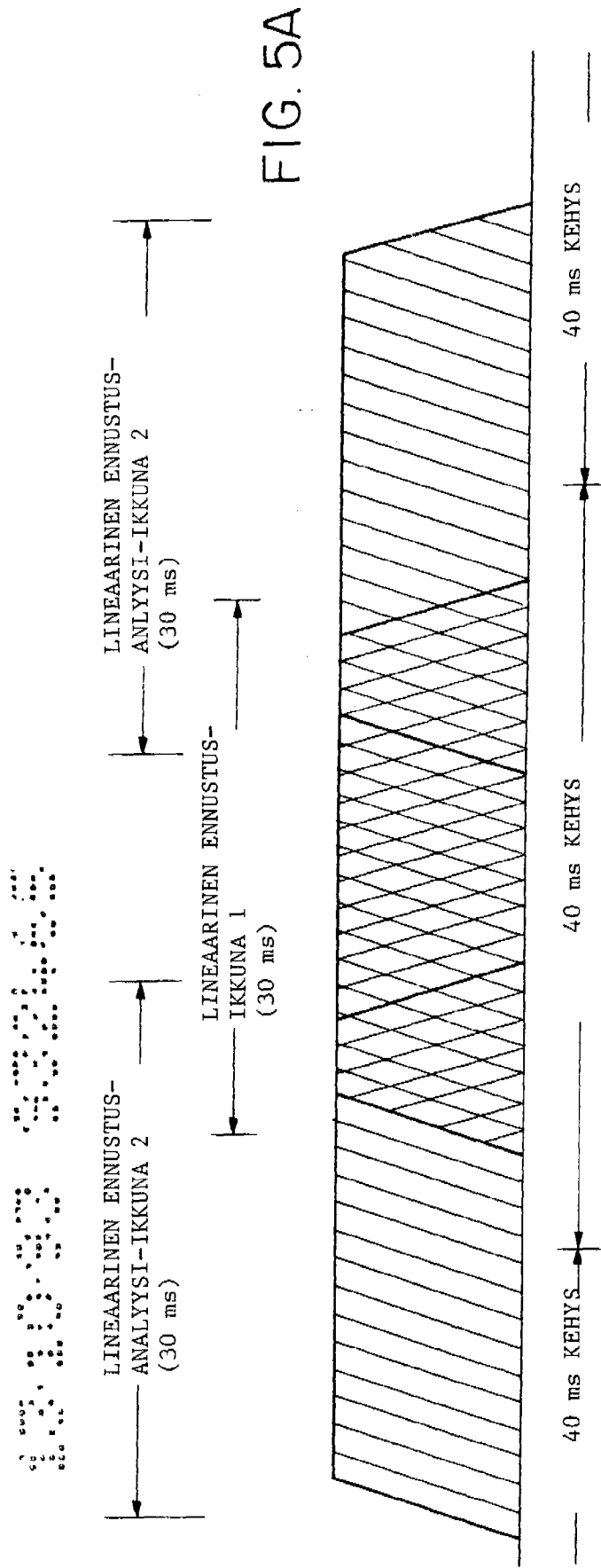
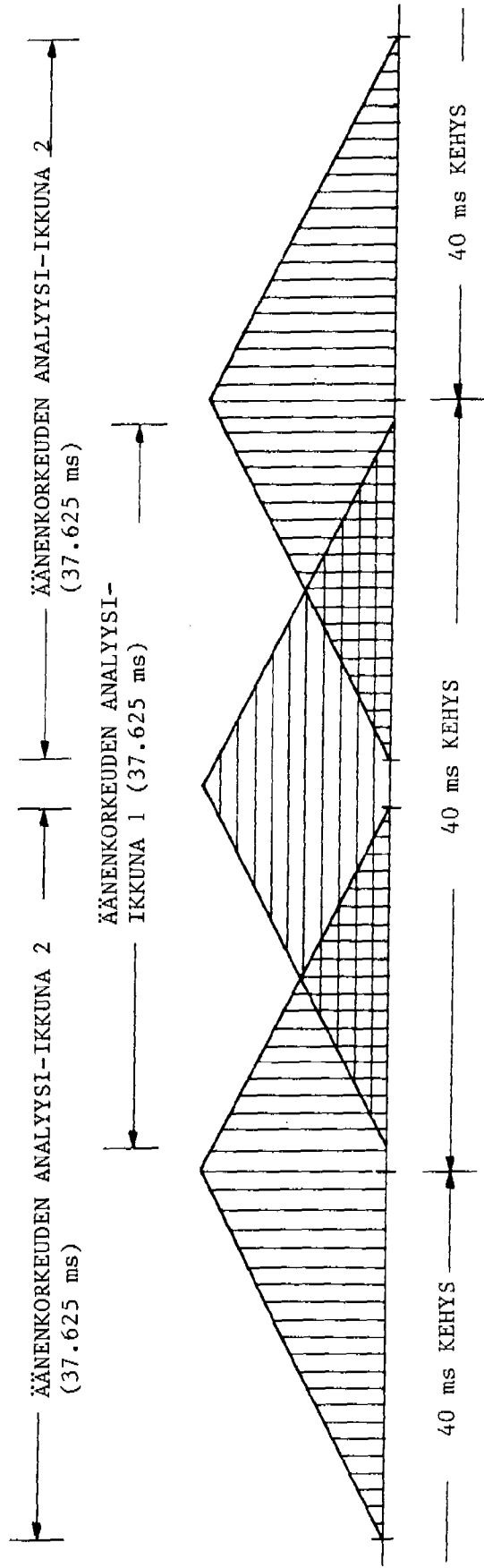


FIG. 5A

FIG. 5B



10 LSF:n Kvantisoimaton Sarja

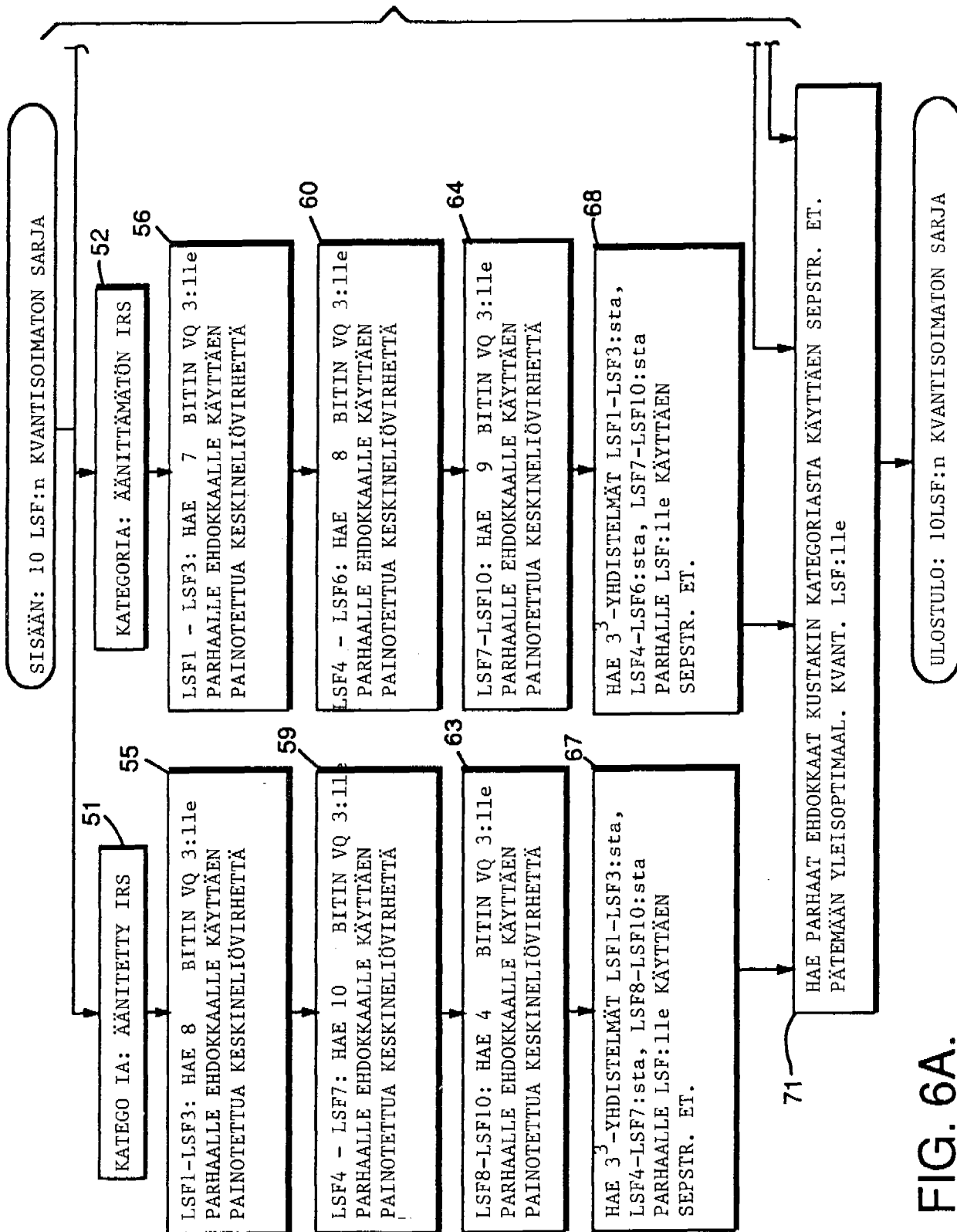


FIG. 6A.

FIG. 6B.:hen

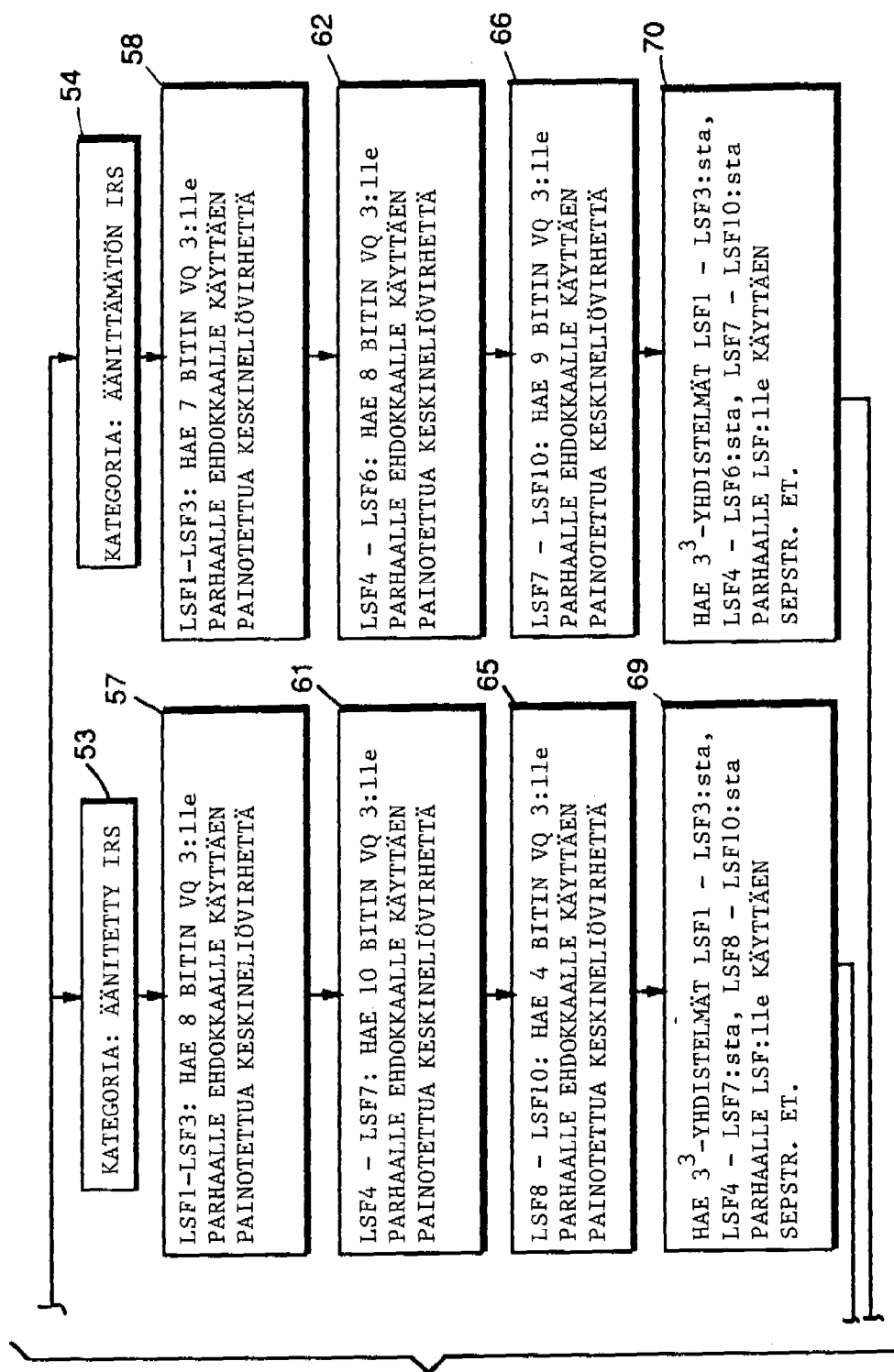
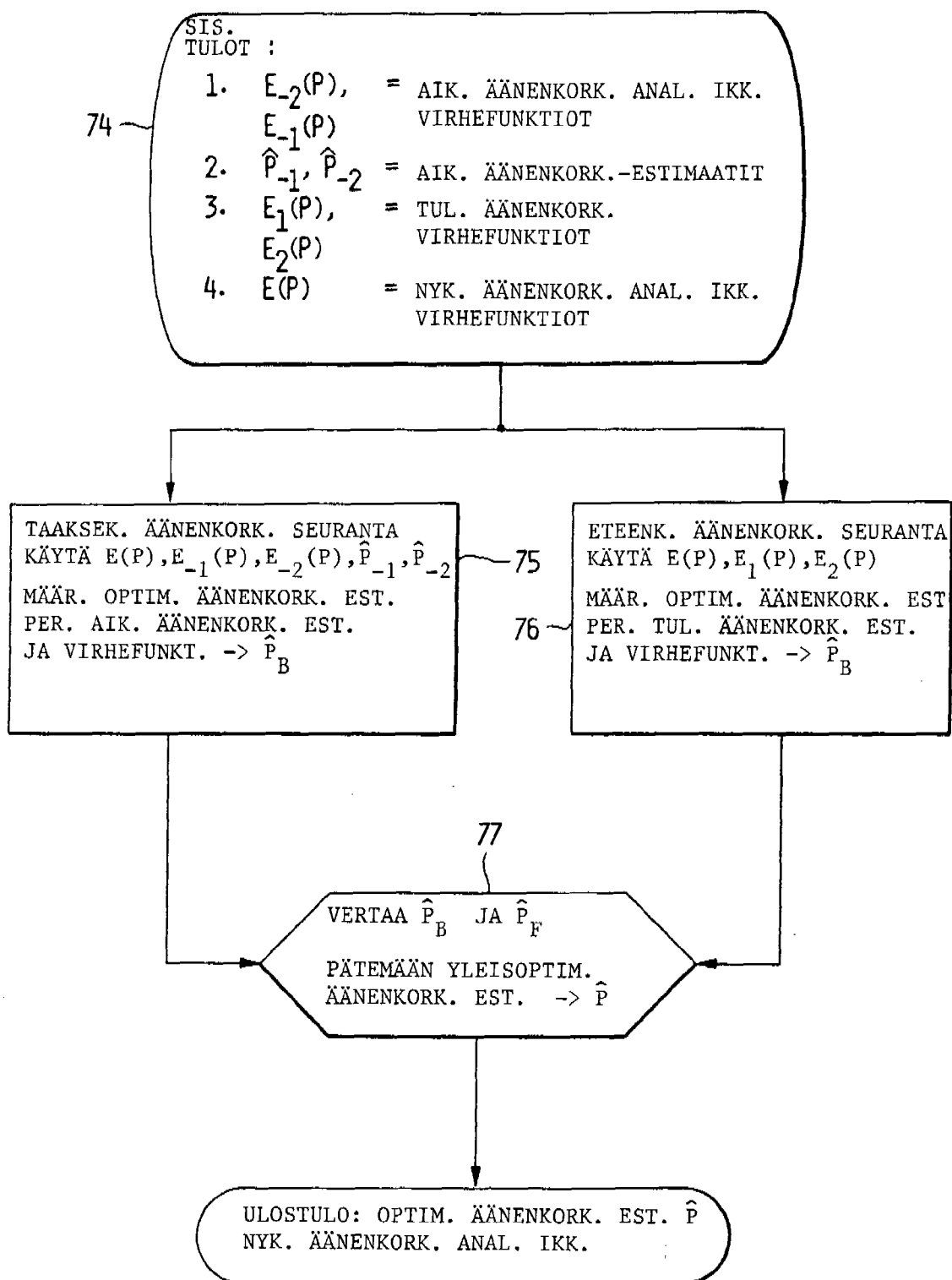


FIG. 6A.

FIG. 6B.

FIG. 7 TUUNNETTU TEKNIikka



101083 800685

FIG. 8

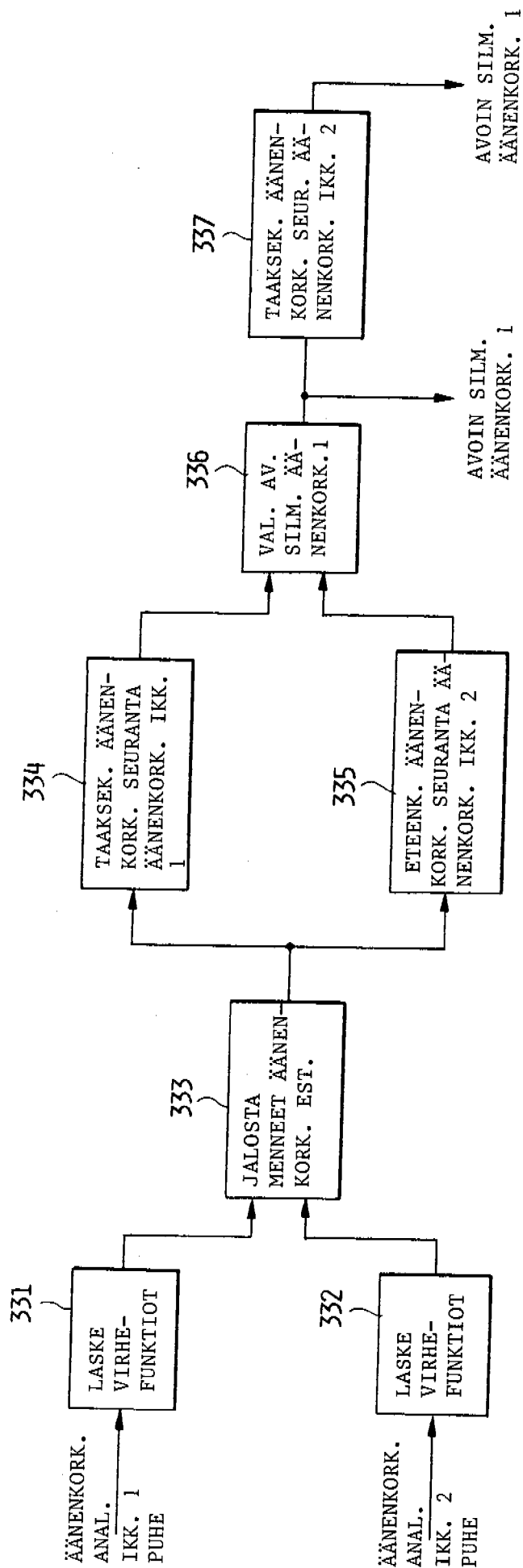


FIG. 9

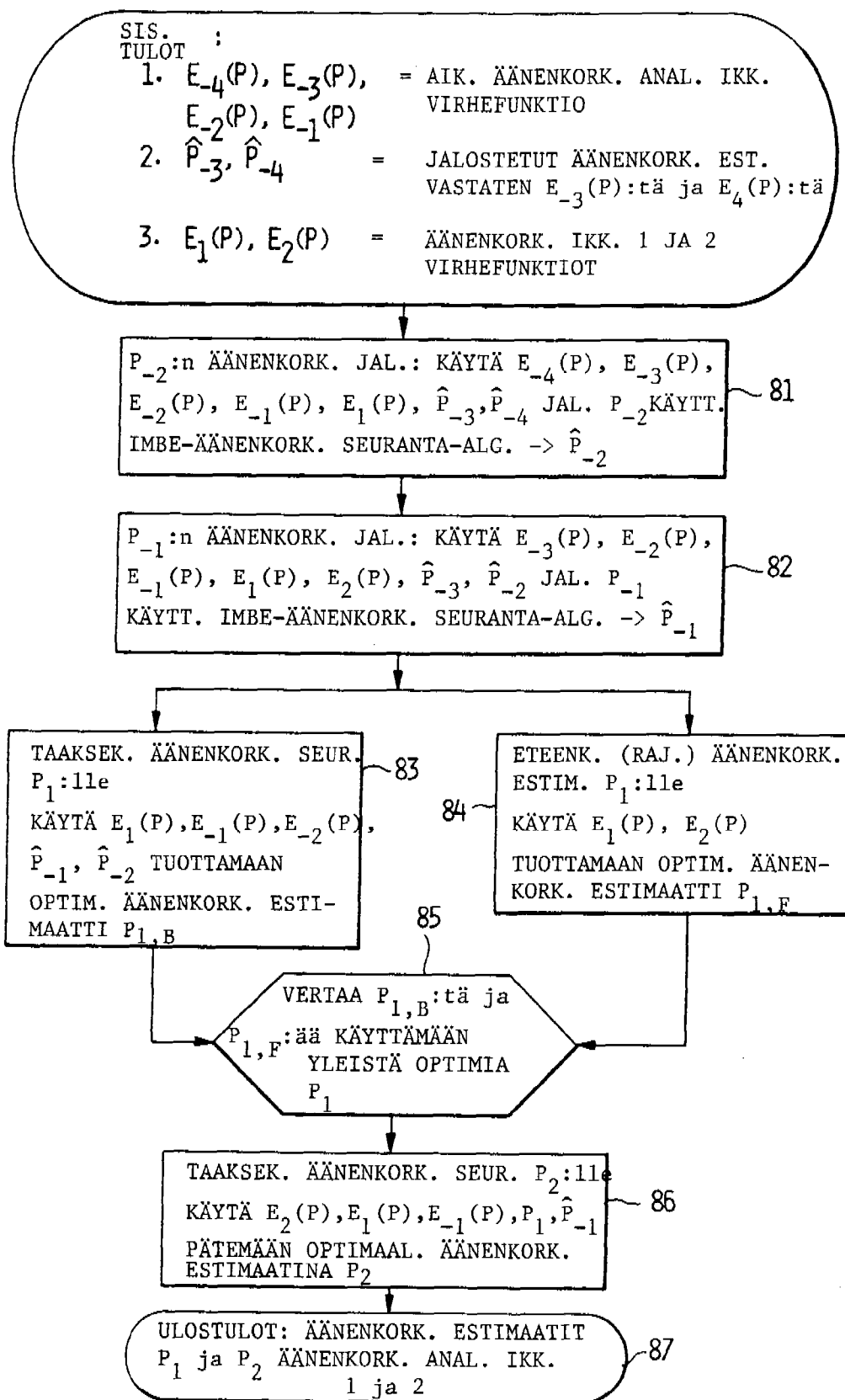


FIG. 10

9/15

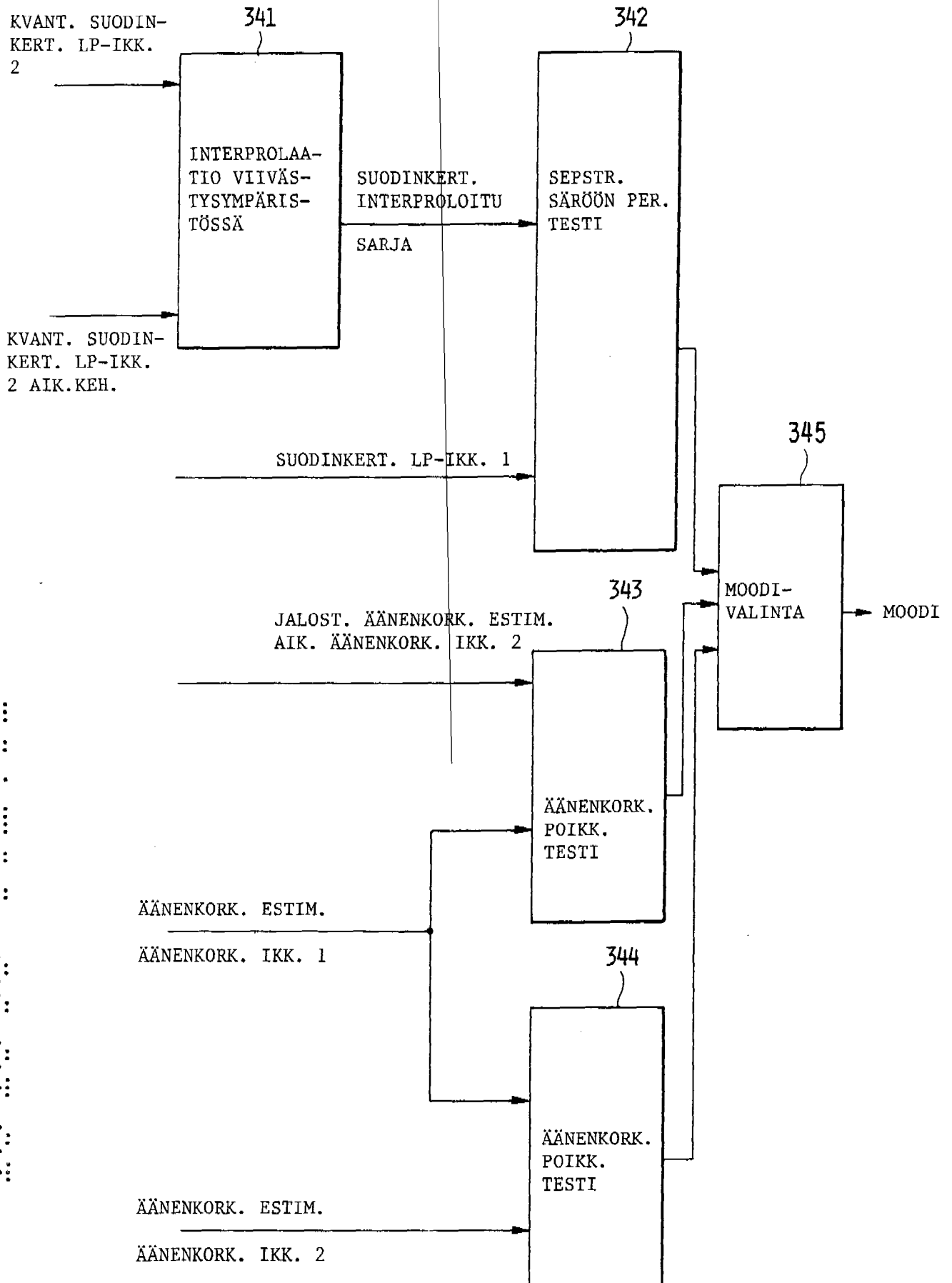


FIG. 11

SIS.TULOT : A_1 = KV.TON SARJA SUODINKERT. 1. LP-ANAL. IKK.

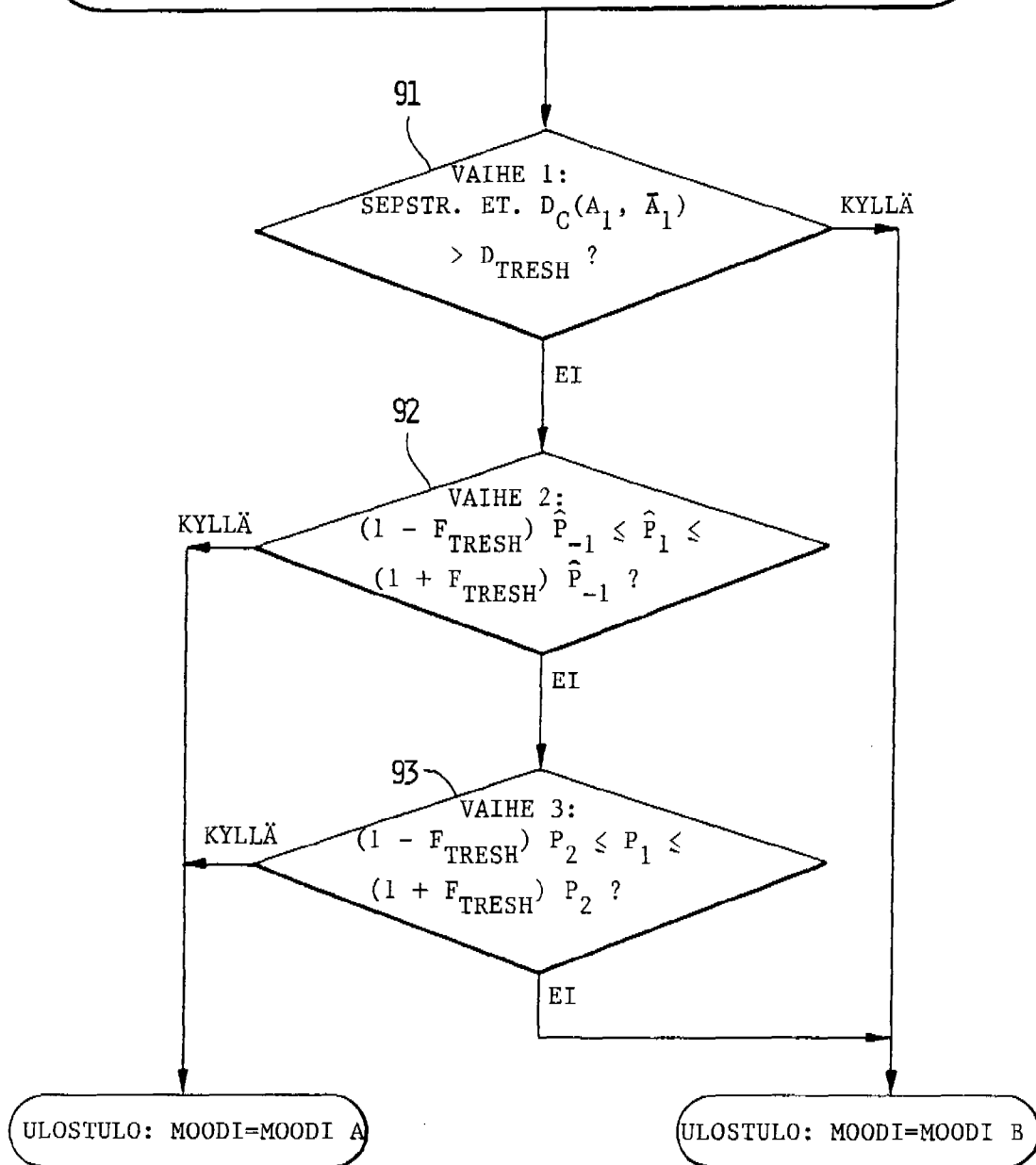
$\bar{A}_1$ = INTERP. SARJA SUODINK. 1. LP-ANAL. IKK.

$\hat{P}_{-1}$ = JAL. AV. SILM. ÄÄNENKORK. EST. AIK. ÄÄNENKORK. IKK.

P_1 = AV. SILM. ÄÄNENKORK. EST. ÄÄNENKORK. IKK. 1

P_2 = AV. SILM. ÄÄNENKORK. EST. ÄÄNENKORK. IKK. 2

D_{TRESH}, F_{TRESH} = KYNNYTEKIJÄT



101080 000000

FIG. 12

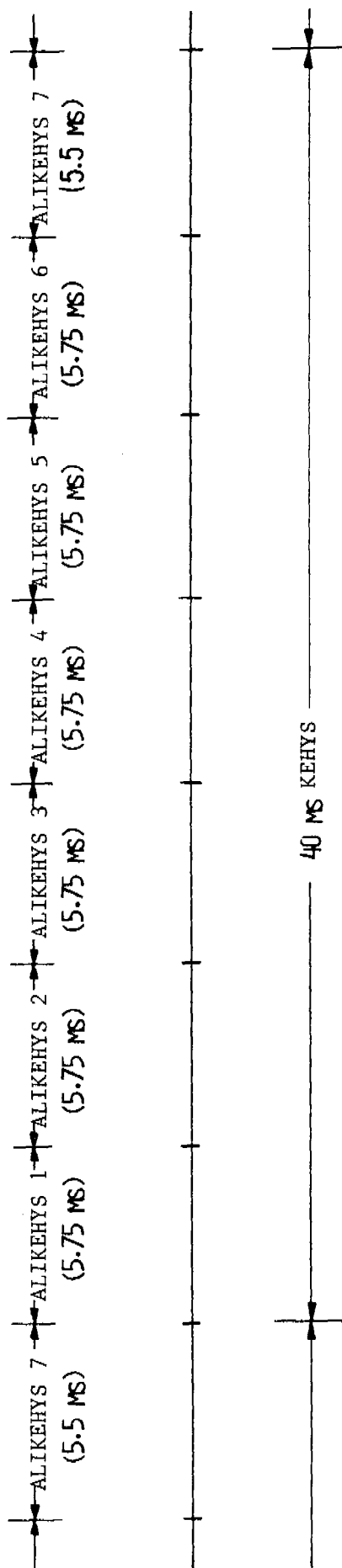
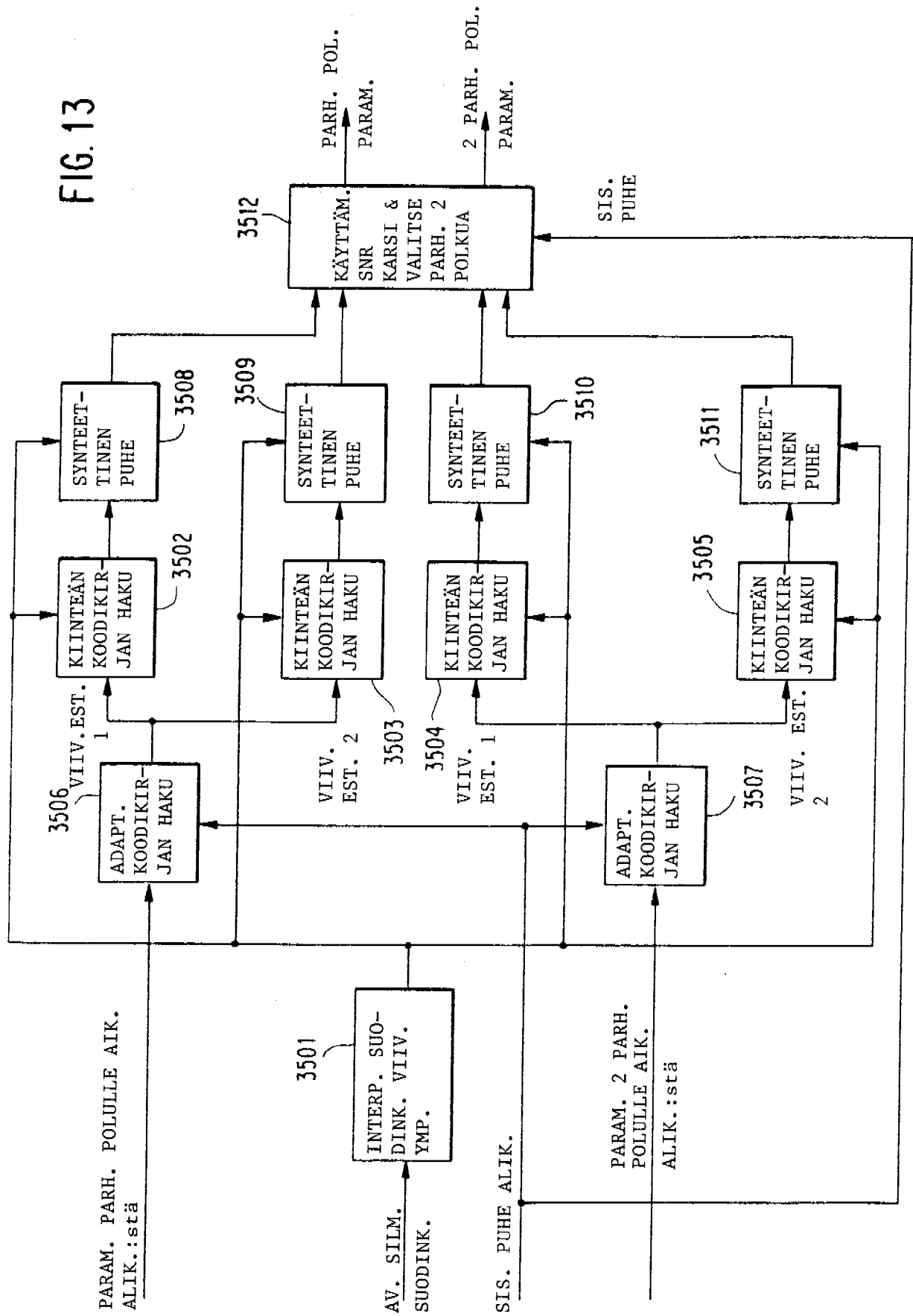
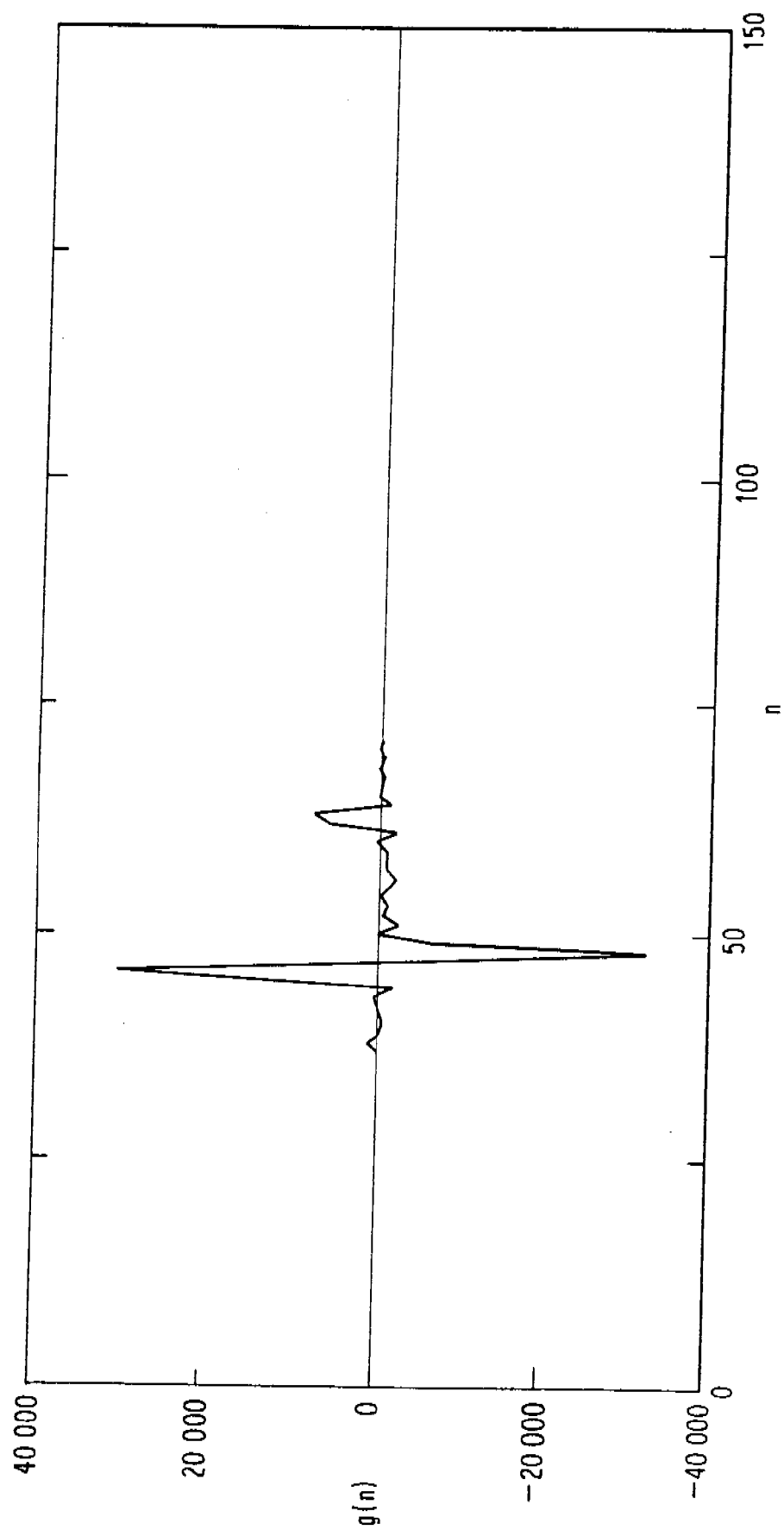


FIG. 13



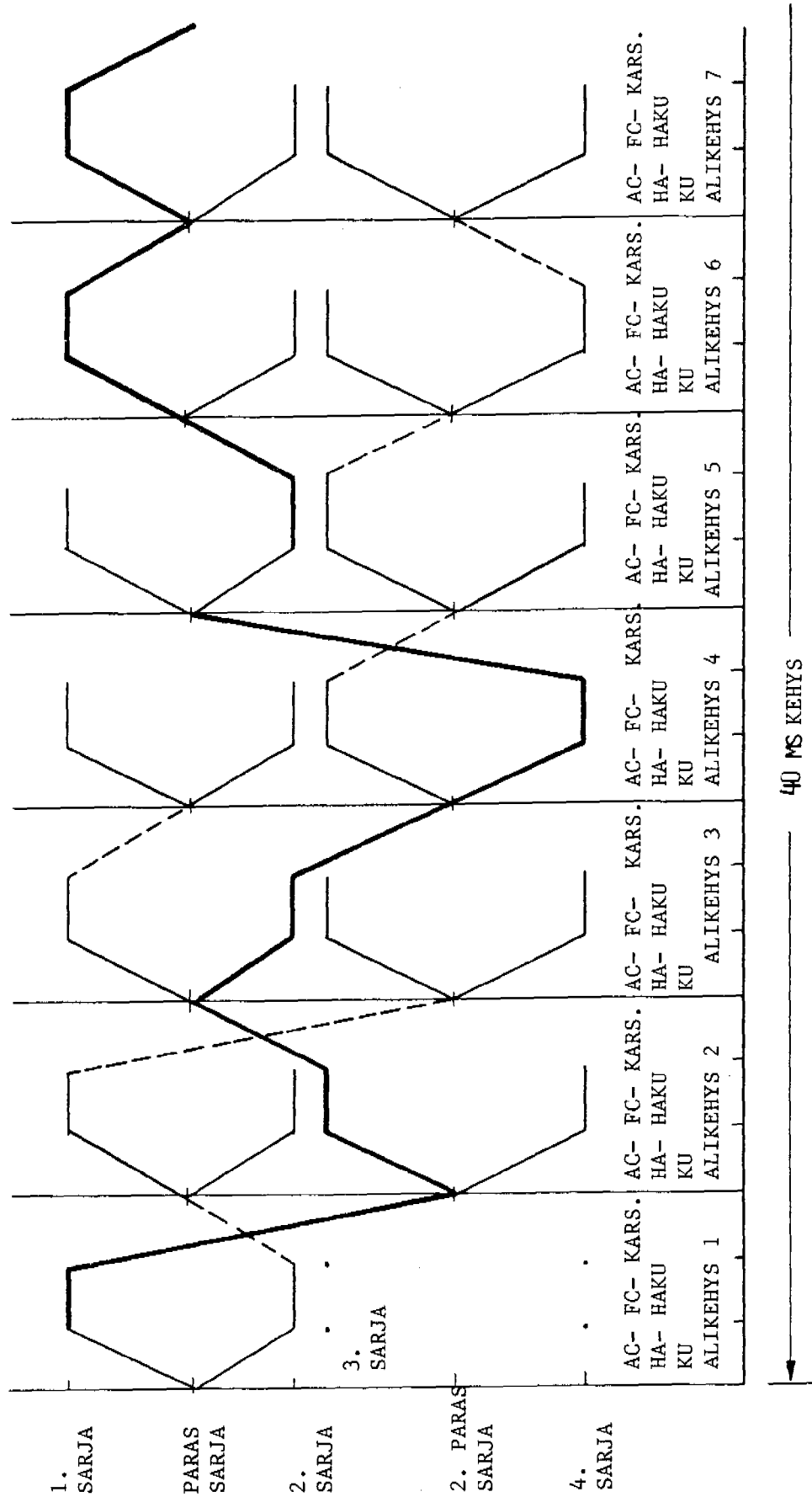
10.10.93 00:04:55

FIG. 14



13.10.83 00.00

FIG. 15



HUOMIOT:

INDEKSIEN/PARAMETRIEN OPTIMISARJA VIIV. PÄÄT. JÄLKEEN: 1-3-2-4-2-1-1; AC: ADAPT. KOODIKIRJA

---: KARTOITUKSET KARS. JÄLK. PARH. JA 2 PARH. SARJAAN: ---: OPTIM. POLKU; FC: KIINTEÄ KOODIKIRJA

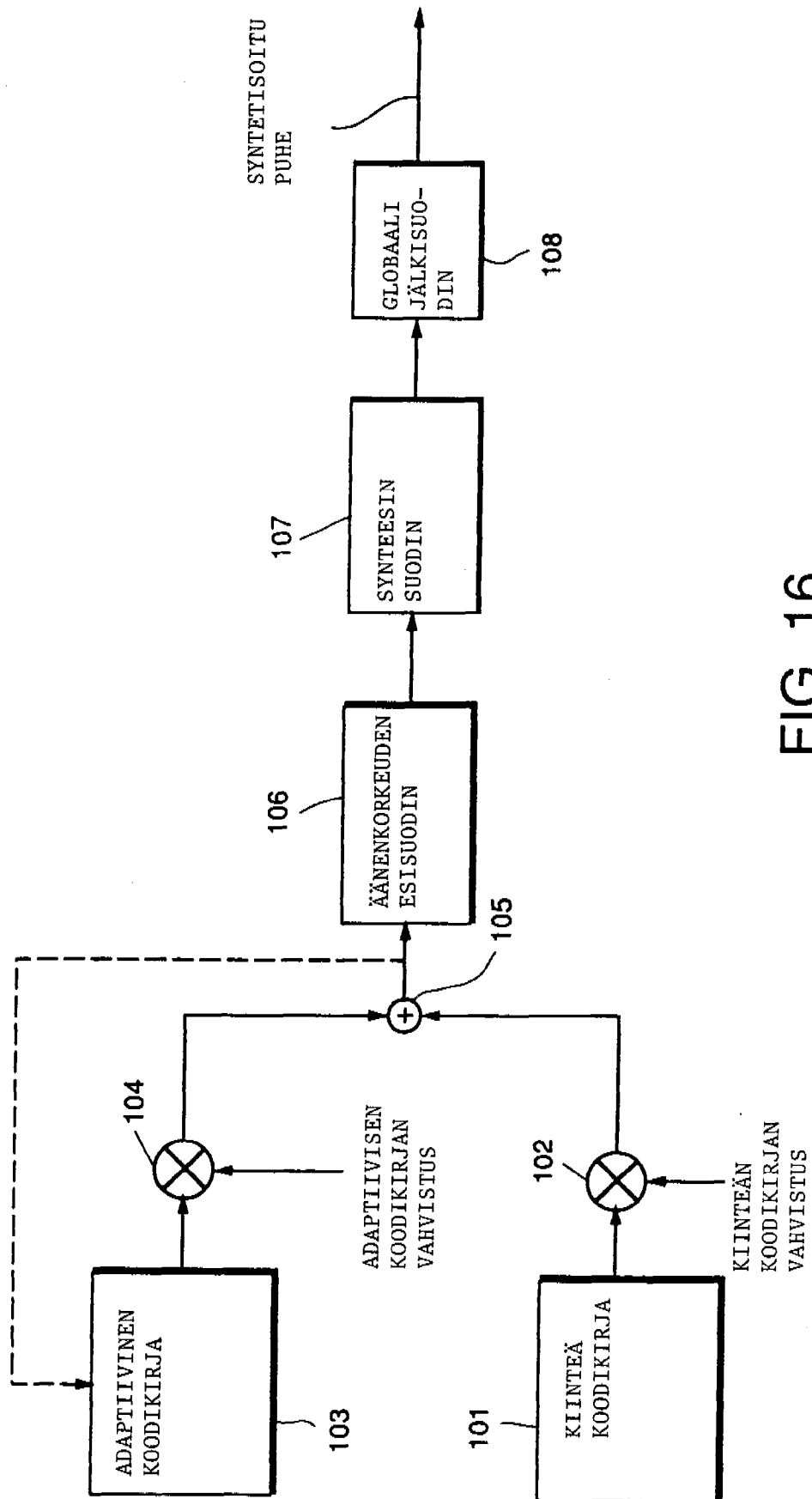


FIG. 16