
Octrooiraad



⑫ A **Terinzagelegging** ⑪ **8603163**

Nederland

⑲ NL

- ⑤4 **Werkwijze en inrichting voor het afleiden van formantfrequenties uit een gedeelte van een spraaksignaal.**
- ⑤1 Int.Cl.: G10L9/04.
- ⑦1 Aanvrager: N.V. Philips' Gloeilampenfabrieken te Eindhoven.
- ⑦4 Gem.: Ir. P.J.P.G. Simons c.s.
Internationaal Octroobureau B.V.
Prof. Holstlaan 6
5656 AA Eindhoven.

-
- ②1 Aanvraag Nr. 8603163.
- ②2 Ingediend 12 december 1986.
- ③2 --
- ③3 --
- ③1 --
- ⑥2 --

-
- ④3 Ter inzage gelegd 1 juli 1988.

De aan dit blad gehechte stukken zijn een afdruk van de oorspronkelijk ingediende beschrijving met conclusie(s) en eventuele tekening(en).

N.V. Philips' Gloeilampenfabrieken te Eindhoven.

Werkwijze en inrichting voor het afleiden van formantfrequenties uit een gedeelte van een spraaksignaal.

De uitvinding heeft betrekking op een werkwijze voor het bepalen van formantfrequenties uit een binnen een zeker tijdinterval liggend gedeelte van een spraaksignaal, waarbij

- voor opeenvolgende tijdstippen liggend binnen het tijdinterval uit het

5 binnen het tijdinterval liggend gedeelte van het spraaksignaal een parameterwaarde wordt afgeleid,

- uit de parameterwaardes wordt een polynoom van een zekere orde bepaald,
- uit de bepaalde polynoom worden de formantfrequenties afgeleid. De uitvinding heeft eveneens betrekking op een inrichting voor het

10 uitvoeren van de werkwijze.

Formanten zijn in werkelijkheid de resonanties van het mondkanaal en worden gekenmerkt door veel energie in het spektrum. Tijdens het praten verandert het mondkanaal voortdurend van vorm en daardoor veranderen ook de formanten wat betreft de ligging op de

15 frekwentie-as en wat betreft de bandbreedte. In een bron-filter-model van spraakproduktie wordt vaak gebruik gemaakt van een beschrijving van het filter in termen van formantfrequenties en -bandbreedten. Ook de spraakanalyse voor de Philips' spraaksynthesechips MEA 8000 en PCF 8200 gebruikt een formantbeschrijving van het spraaksignaal, zie (1) en

20 (2).

De redenen waarom men een formantbeschrijving wil toepassen zijn:

- er is een zuinige codering mogelijk,
- men heeft te maken met fysisch te interpreteren gegevens, waardoor manipulaties inzichtelijk zijn, zoals bijvoorbeeld konkatenering van

25 difoonsegmenten en editing ten behoeve van de spraaksynthesechip.

Hierboven is de indruk gewekt als zou het spraaksignaal altijd door een aantal formanten (= resonanties) beschreven kunnen worden. In dat geval bestaat het filter in het bron-filter-model uit alleen resonanties (all pole filter). In lopende spraak voldoet het

30 spraakproduktiesysteem niet altijd aan dit model: er zijn klanken waarvoor het model minder formanten zou moeten bevatten of naast formanten ook nulpunten (= antiresonanties: dit is een frekwentiegebied

8603163

waar een aan resonantie tegengesteld verschijnsel werkzaam is, waardoor het signaal niet wordt opgeslingerd maar als het ware wordt weggezogen en waar lokaal weinig energie in het spektrum is). In een praktisch systeem echter is de structuur van het bron-filter-model vastgelegd en
5 daardoor ook het aantal formanten. Dit niet aangepast zijn van het gehanteerde model aan alle werkelijk optredende situaties maakt dat bij de spraaksynthese een operationele definitie aan de formanten wordt gegeven. Het spraaksynthesefilter bevat slechts een vast aantal formanten (en geen nulpunten) en de daarbij behorende spraakanalyse
10 heeft de opdracht de modelparameters te vinden onafhankelijk van het feit of het model past bij de spraakproductie.

Een formantanalyse is uitvoerig beschreven in (3).

Bij deze formantanalyse doen zich een tweetal problemen voor:

- niet altijd wordt het voorgeschreven aantal formanten gevonden,
- 15 - af en toe faalt de analyse om numerieke redenen: het toegepaste algoritme convergeert niet.

De uitvinding beoogt nu een werkwijze en een inrichting voor het uitvoeren van de werkwijze aan te geven die in staat is het voorgeschreven aantal operationeel gedefinieerde formanten in alle
20 gevallen te bepalen waarbij gebruik gemaakt wordt van een algoritme dat in alle gevallen convergeert.

De werkwijze volgens de uitvinding heeft daartoe het kenmerk, dat een Split Levinson algoritme wordt uitgevoerd, waarbij in een aantal achtereenvolgende rekursiestappen uit de parameterwaardes
25 telkens een singuliere prediktorpolynoom van een hogere orde wordt bepaald, en dat na de laatste rekursiestap uit de in de laatste rekursiestap verkregen singuliere prediktorpolynoom de formantfrequenties worden afgeleid. De werkwijze kan daarbij verder zijn gekenmerkt doordat in een rekursiestap de nulpunten van de in deze
30 rekursiestap bepaalde singuliere prediktorpolynoom worden afgeleid, daarbij gebruik makend van de nulpunten berekend tijdens de voorgaande rekursiestap, en dat na de laatste rekursiestap uit de in deze rekursiestap verkregen nulpunten de formantfrequenties worden afgeleid. Het bepalen van de nulpunten van de singuliere prediktorpolynomen gaat
35 eenvoudiger dan het bepalen van de nulpunten volgens de bekende methode. De nulpunten van de polynoom die volgens de bekende werkwijze zijn verkregen liggen binnen de eenheidscirkel terwijl de nulpunten van

een singuliere prediktorpolynoom op de eenheidscirkel liggen. Dit heeft bovendien tot gevolg dat de nulpunten op eenvoudiger wijze zijn te berekenen en dat bovendien altijd voldoende nulpunten worden gevonden, zodat in feite een robuuste methode voor het bepalen van

5 formantfrequenties is verkregen.

De werkwijze kan verder zijn gekenmerkt, doordat bij elk der aldus gevonden formantfrequenties de bijbehorende bandbreedte wordt bepaald uitgaande van de parameterwaardes en de berekende formantfrequenties, door middel van een minimaliseringsalgoritme. Men

10 heeft daarmee weer alle grootheden afgeleid benodigd om synthetische spraak te genereren, zoals reeds gedaan wordt met de eerder genoemde spraakchips MEA 8000 en PCF 8200.

De inrichting voor het uitvoeren van de werkwijze, voorzien van

- 15 - een ingangsklem voor het ontvangen van een spraaksignaal,
- een eerste eenheid voor het, voor opeenvolgende tijdstippen liggend binnen een tijdinterval, uit het binnen dit tijdinterval gelegen gedeelte van het spraaksignaal afleiden van een parameterwaarde, met een ingang gekoppeld met de ingangsklem, en een uitgang,
- 20 - een tweede eenheid voor het bepalen van een polynoom van een zekere orde uit de parameterwaardes, met een ingang gekoppeld met de uitgang van de eerste eenheid, en een uitgang, en
- een derde eenheid voor het afleiden van de formantfrequenties uit de bepaalde polynoom, met een ingang gekoppeld met de uitgang van de tweede
- 25 eenheid en een uitgang voor het afgeven van de formantfrequenties, heeft het kenmerk, dat de tweede eenheid is ingericht voor het uitvoeren van een Split Levinson algoritme, en voor het daarin in een aantal achtereenvolgende rekursiestappen uit de parameterwaardes afleiden van telkens een singuliere prediktorpolynoom van een hogere orde en dat de
- 30 derde eenheid is ingericht voor het afleiden van de formantfrequenties uit de in de laatste rekursiestap verkregen singuliere prediktorpolynoom.

De tweede eenheid kan daarbij verder zijn ingericht voor het in een rekursiestap afleiden van de nulpunten van de in deze

35 rekursiestap bepaalde singuliere prediktorpolynoom, daarbij gebruik makend van de nulpunten berekend tijdens de voorgaande rekursiestap, en dat de derde eenheid ingericht is voor het afleiden van de

formantfrequenties uit de nulpunten verkregen in de laatste rekursiestap. Wil men behalve de op bovenstaande wijze verkregen formantfrequenties ook nog de bandbreedtes bepalen dan kan daartoe de derde eenheid verder zijn ingericht voor het bij elk der aldus gevonden
 5 formantfrequenties bepalen van de bijbehorende bandbreedte, uitgaande van de parameterwaardes en de berekende formantfrequenties, door middel van een minimaliseringsalgoritme.

De uitvinding zal hierna in de figuurbeschrijving nader worden uiteengezet. Hierin toont

10 figuur 1 nulpunten van het A-filter uit de LPC-analyse, liggend binnen de eenheidscirkel en nulpunten van de singuliere prediktorpolynoom, liggend op de eenheidscirkel,

 figuren 2 en 3 het gedrag van de nulpunten verkregen voor opeenvolgende rekursiestappen in het Split Levinson algoritme,

15 figuur 4 een stroomdiagram van de werkwijze,

 figuur 5 een stroomdiagram van het programmagedeelte waarin het Split Levinson algoritme wordt toegepast,

 figuur 6 een inrichting voor het uitvoeren van de werkwijze.

20 De formantbepaling in de bekende werkwijze geschiedt door het berekenen van een all pole filter met behulp van de LPC-analyse, dat vervolgens wordt ontleedt in tweede orde sekties. De LPC-analyse is een uit de literatuur bekende analysemethode, zie bijvoorbeeld (5). Bij de LPC-analyse neemt men een stuk signaal van ongeveer 25 ms en
 25 vermenigvuldigt dat met een Hamming venster en berekent de autokorrelatie-koëfficiënten. Met behulp van het zogenaamde Levinson algoritme wordt nu een polynoom $A(z)$ ($1/A(z)$ = het all pole filter) van een zekere orde bepaald. Dit is een rekursief algoritme waarin voor elke rekursie-slag een A-polynoom wordt berekend waarvan de nulpunten binnen
 30 de eenheidscirkel liggen.

Achtereenvolgens:

$$A_0(z) = 1 \qquad (1.1)$$

35 $A_1(z) = 1 + a_{1,1}z^{-1} \qquad (1.2)$

$$A_2(z) = 1 + a_{2.1}z^{-1} + a_{2.2}z^{-2} \quad (1.3)$$

$$A_m(z) = 1 + a_{m.1}z^{-1} + \dots + a_{m.m}z^{-m} \quad (1.4)$$

5

Bij elke rekursie verandert het A-polynoom in zijn geheel. Dat de nulpunten steeds binnen de eenheidscirkel liggen garandeert een stabiel synthesefilter en is een gevolg van het toepassen van de autokorrelatiemethode. De nulpunten van dit polynoom zijn toegevoegde

10 complexe nulpuntparen, òf reële nulpunten, zie figuur 1. In figuur 1 geven de open rondjes de toegevoegde complexe nulpunt-paren aan en geven de gesloten rondjes de reële nulpunten aan. De nulpunt-paren (inklusief reële) zijn te schrijven als:

$$15 \quad N(z) = 1 + pz^{-1} + qz^{-2} \quad (2)$$

Als men het A-polynoom $A(z)$ schrijft als:

$$A(z) = 1 + a_1z^{-1} + \dots + a_mz^{-m} \quad (3)$$

20

kan dat ontleedt worden in tweede orde sekties:

$$25 \quad A(z) = \prod_{j=1}^{M/2} N_j(z^{-1}) = \prod_{j=1}^{M/2} (1 + p_j z^{-1} + q_j z^{-2})$$

Het afsplitsen van deze (p_j, q_j) -paren kan geschieden met behulp van het zogenaamde bairstow-algoritme, dat uit de handboeken bekend is, zie onder andere (6).

30 Toegevoegde complexe nulpunt-paren vertegenwoordigen een resonantie (= formant) en de p_j, q_j getallen geven de formantfrequentie en -bandbreedte als volgt:

$$p_j = -2 \cdot \exp(-\pi B_j T) \cdot \cos(2\pi F_j T) \quad (5a)$$

35

$$q_j = \exp(-2\pi B_j T) \quad (5b)$$

waarin $T = 1/F_s$ de bemonsteringsperiode; hieruit zijn B_j en F_j te bepalen.

Reële nulpunten kunnen niet naar formantgegevens omgerekend worden omdat deze geen resonantie beschrijven, maar veel meer
5 het spektrum een zekere helling geven.

De twee in de inleiding genoemde problemen bij de huidige formantbepaling kan men nu beter formuleren:

- de aanwezigheid van reële nulpunten van het A-polynoom, waardoor geen formantfrequentie en -bandbreedte bepaald kan worden,
- 10 - het af en toe falen van het bairstow-algoritme om numerieke redenen die niet echt bekend zijn. Het algoritme blijft dan itereren zonder te konvergeren.

Het zogenaamde Split Levinson algoritme is ontwikkeld door Genin en Delsarte (4), en een van de eigenschappen is, dat ongeveer
15 de helft van het aantal vermenigvuldigingen nodig is om een LPC-analyse uit te voeren vergeleken met het klassieke Levinson algoritme. Dit is mogelijk doordat in plaats van de A-polynomen nu de zogenaamde singuliere prediktor polynomen worden gebruikt. Deze zijn symmetrisch en daardoor liggen de nulpunten op de eenheidscirkel en deze polynomen
20 bestaan dus globaal gesproken uit half zo veel betekenisvolle coëfficiënten.

Het aantrekkelijke van dit algoritme ligt in de eigenschappen van de singuliere prediktor polynomen (SPP). De SPP worden gedefinieerd door

$$25 \quad P_k(z) = A_{k-1}(z) + z^{-1} \hat{A}_{k-1}(z) \quad (6)$$

waarbij $A_k(z)$ het A-polynoom bij de k-de rekursie van het normale Levinson algoritme is en voor $\hat{A}_k(z)$ geldt:

$$30 \quad \hat{A}_k(z) = z^{-k} \cdot A_k(z^{-1}) \quad (7)$$

$\hat{A}_k(z)$ is de reciproke polynoom van $A_k(z)$.

Deze SPP zijn zoals gezegd symmetrische polynomen en ze
35 hebben daarom nulpunten die op de eenheidscirkel liggen en niet erbinnen zoals bij de $A_k(z)$ het geval is.

Deze SPP zijn ook verwant aan de polynomen die een rol

spelen bij de LSP-analyse (Line Spectrum Pairs) (7). Op grond van de definitie en de eigenschappen van $A_k(z)$ kan men een rekurrente betrekking afleiden voor de SPP:

$$5 \quad P_k(z) = (1+z^{-1}) P_{k-1}(z) - \alpha_{k-1} z^{-1} P_{k-2}(z) \quad (8)$$

hier is α_{k-1} een getal dat wordt berekend uit de gegeven autokorrelatiecoëfficiënten.

Het is bekend (7) dat de positie van de nulpunten op de
 10 eenheidscirkel van deze SPP voor even waarden van de orde in de buurt is van de formantposities zoals men die afleidt uit het A-polynoom. Deze overeenkomst is des te beter naarmate de pool dichter bij de eenheidscirkel ligt of met andere woorden de bandbreedte van de formant
 15 afgeleid uit de posities van de nulpunten van de singuliere prediktor polynoom op de eenheidscirkel. Het probleem is nu vereenvoudigd van het vinden van de nulpunten van het A-polynoom, die overal binnen de eenheidscirkel kunnen liggen, tot het vinden van de nulpunten van de
 20 met een kruis aangegeven punten op de eenheidscirkel in figuur 1. Het vinden van deze nulpunten van de singuliere prediktor polynoom wordt verder nog vergemakkelijkt, doordat de nulpunten in de suksessievelijke rekursiestappen heel systematisch verschuiven.

De rekursiestappen worden op de volgende wijze
 25 doorlopen. In de eerste rekursiestap wordt $P_0(z) = 1$ genomen. In de tweede rekursiestap wordt $P_1(z) = 1+z^{-1}$ genomen. Dit volgt rechtstreeks uit de formules (1.1), (6) en (7). Het nulpunt $np_{1,1}$ van deze polynoom ligt bij $z^{-1} = -1$ ofwel $w = \pi$, waarbij w het argument is van het (komplexe) nulpunt. In de derde rekursiestap wordt $P_2(z)$
 30 berekend gebruik makend van de formule (8):

$$P_k(z) = (1+z^{-1}) P_{k-1}(z) - \alpha_{k-1} z^{-1} P_{k-2}(z) \quad (8)$$

waarbij

$$35 \quad \alpha_{k-1} = \tau_{k-1} / \tau_{k-2} \quad (9)$$

$$\tilde{\tau}_{k-1} = (r_0 + r_{k-1}) + (r_1 + r_{k-2}) p_{k-1.1} + \dots \quad (10)$$

en $p_{k.i}$ volgt uit de algemene formule voor $P_k(z)$ te weten

$$5 \quad P_k(z) = 1 + p_{k.1}z^{-1} + p_{k.2}z^{-2} + \dots + p_{k.1}z^{-k+1} + z^{-k} \quad (11)$$

ofwel

$$P_2(z) = 1 + p_{2.1}z^{-1} + z^{-2}$$

10

Voor het berekenen van $P_2(z)$ geldt dus dat

$$P_{k-1.1} = P_{k-1.2} = \dots = 0$$

15 en dus is

$$\tilde{\tau}_1 = r_0 + r_1$$

Bovendien kiest men: $\tilde{\tau}_0 = r_0/2$

20

Bijgevolg wordt $P_2(z)$:

$$\begin{aligned} P_2(z) &= (1+z^{-1})^2 - \alpha_1 z^{-1} \\ &= 1 + (2-\alpha_1) z^{-1} + z^{-2} \\ 25 \quad &= z^{-1} \{(2-\alpha_1) + (z^{-1}+z)\} \end{aligned}$$

Substitueren we hierin $z = e^{jw}$, hetgeen betekent dat $z+z^{-1} = 2\cos w$ dan krijgen we:

$$30 \quad P_2(z) = e^{-jw} \{(2-\alpha_1) + 2\cos w\}$$

De tweedegraads polynoom $P_2(z)$ is nu teruggebracht tot een eerstegraads polynoom met nulpunten op het interval $(-1,+1)$ in plaats van op de eenheidscirkel.

35 We vinden een nulpunt $np_{2.1}$ dat ligt in het interval bepaald door $np_{1.1}$ ($= -1$) en $+1$, zie figuur 2.

Vervolgens wordt in de vierde rekursiestap $P_3(z)$

berekend gebruik makend van de formules (8), (9), (10) en (11). Men vindt een vergelijking van de vorm:

$$\begin{aligned}
 P_3(z) &= 1 + p_{3.1}z^{-1} + p_{3.1}z^{-2} + z^{-3} \\
 5 \quad &= (1+z^{-1}) [1+(p_{3.1}-1)z^{-1} + z^{-2}]
 \end{aligned}$$

Deze vergelijking is te delen door $1+z^{-1}$, hetgeen een nulpunt $np_{3.1}$ oplevert bij $z^{-1} = -1$ ofwel $w = \pi$.

Wat overblijft is weer een tweedegraadsvergelijking die weer omgewerkt
 10 kan worden op de wijze zoals aan de hand van $P_2(z)$ beschreven.
 Men vindt dan een nulpunt $np_{3.2}$ dat ligt in het interval bepaald door $np_{2.1}$ en $+1$, zie figuur 2.

Vervolgens wordt in de vijfde rekursiestap $P_4(z)$
 berekend gebruik makend van de formules (8), (9), (10) en (11):

15

$$\begin{aligned}
 P_4(z) &= 1 + p_{4.1}z^{-1} + p_{4.2}z^{-2} + p_{4.1}z^{-3} + z^{-4} \\
 &= z^{-2} (z^2 + p_{4.1}z + p_{4.2} + p_{4.1}z^{-1} + z^{-2}) \\
 &= z^{-2} [(z^2+z^{-2}) + p_{4.1}(z+z^{-1}) + p_{4.2}]
 \end{aligned}$$

20 Substitueren we weer $z = e^{jw}$, dan is

$$z + z^{-1} = 2 \cos w \text{ en } z^k + z^{-k} = 2 \cos kw$$

$$P_4(z) = e^{-2jw} [2 \cos 2w + 2p_{4.1} \cos w + p_{4.2}]$$

25

En dit is altijd te schrijven in machten van $y = \cos w$; in dit geval met $\cos 2w = 2 \cos^2 w - 1$.

$$P_4(z) = e^{-2jw} [4y^2 + 2p_{4.1}y + (p_{4.2}-2)]$$

30

De vierdegraads polynoom $P_4(z)$ is nu teruggebracht tot een tweedegraads polynoom met nulpunten weer op het interval $(-1,+1)$ in plaats van op de eenheidscirkel. In het bijzonder ligt er een nulpunt $np_{4.1}$ tussen $np_{3.1}$ en $np_{3.2}$ in en ligt er een nulpunt $np_{4.2}$

35 tussen $np_{3.2}$ en $+1$ in, zie figuur 2.

Samenvattend:

In het Split Levinson algoritme zien de SPP in de achtereenvolgende rekursiestappen eruit als:

$$\begin{aligned}
 5 \quad k = 0 & \quad P_0(z) = 1 \\
 k = 1 & \quad P_1(z) = 1 + z^{-1} \\
 k = 2 & \quad P_2(z) = 1 + p_{2.1}z^{-1} + z^{-2} \\
 k = 3 & \quad P_3(z) = 1 + p_{3.1}z^{-1} + p_{3.1}z^{-2} + z^{-3} \\
 & \quad \quad \quad = (1 + z^{-1})(1 + (p_{3.1}-1)z^{-1} + z^{-2}) \\
 10 \quad k = 4 & \quad P_4(z) = 1 + p_{4.1}z^{-1} + p_{4.2}z^{-2} + p_{4.1}z^{-3} + z^{-4} \\
 & \quad \quad \quad \text{enzovoorts.}
 \end{aligned}$$

Het is een eigenschap van deze SPP $P_k(z)$ dat de nulpunten van $P_k(z)$ liggen in een interval dat is af te leiden uit de nulpunten van $P_{k-1}(z)$. Zie figuur 2: voor $k = 1$ is het nulpunt $np_{1.1}$ 15 $= -1$, voor $k = 2$ ligt het nulpunt in het interval $(np_{1.1}, +1)$. Voor $k = 3$ is een nulpunt $np_{3.1} = -1$ en het ander nulpunt $np_{3.2}$ ligt in het interval $(np_{2.1}, +1)$, enzovoorts.

Het vinden van een nulpunt in een interval waarvan bekend is dat er slechts een aanwezig is leidt altijd tot succes. In het 20 algoritme worden nu vanaf het begin (vanaf $k = 3$) steeds de posities van de nulpunten bepaald, zie ook figuur 3.

De formantfrequenties worden op de volgende wijze berekend uit de nulpunten bepaald in de laatste rekursiestap. Daar een nulpunt $np_{i,j}$ de lengte van de projectie op de horizontale as (zie 25 figuur 1) van de eenheidsvektor naar een bepaald punt op de eenheidscirkel aangeeft geldt:

$$np_{i,j} = \cos(2\pi f_j T)$$

30 waarbij $T = 1/f_s$, de bemonsteringsperiode en f_s de bemonsteringsfrequentie.

Hieruit volgt dat de formantfrequentie

$$35 \quad f_j = f_s \arccos(np_{i,j})/2\pi$$

waarbij j loopt van 1 tot en met $1/2 M$ en i gelijk is aan M .

8003163

Het getal M wordt bepaald door het aantal formanten dat men verwacht binnen het te analyseren frekwentiebereik. Is de bandbreedte van het te analyseren frekwentiebereik bij voorbeeld 5000 Hz dan bevinden zich daarbinnen voor een mannenstem vijf en voor een vrouwenstem vier
 5 formanten. M is in dat geval 10 respektievelijk 8. Is de bandbreedte bijvoorbeeld 8000 Hz dan bevinden zich binnen dit frekwentiegebied 8 formanten voor een mannenstem en 6 formanten voor een vrouwenstem. M is nu 16 respektievelijk 12. Duidelijk mag zijn dat M dus gelijk genomen wordt aan twee maal het verwachte aantal formanten binnen het
 10 frekwentiegebied.

De bandbreedte-informatie bij de aldus gevonden formantfrequenties moet nu nog bepaald worden. Dit probleem wordt opgelost door een minimaliseringstechniek toe te passen, met de bandbreedten als onbekenden. Hiertoe doet men uit de tabel van mogelijke
 15 bandbreedten een keuze voor elke formant. Daaruit is een A-polynoom te berekenen, waarvan men kan nagaan hoe goed dit past bij het binnenkomende signaal. Dus kan men ook berekenen welke keuze uit de tabel de beste passing heeft met het binnenkomende signaal. De passing tussen een a-filter en het binnenkomende signaal kan nu bepaald worden
 20 met behulp van de (al berekende) autokorrelatiecoëfficiënten. Stel dat $\bar{A}(z^{-1})$ het \bar{a} -filter is dat tot stand is gekomen door voor alle nog onbekende bandbreedten een waarde uit de beschikbare tabel te kiezen. Dan is de gemaakte fout

$$\begin{aligned}
 25 \quad E &= \sum_n (s_n + \sum_{k=1}^M \bar{a}_k s_{n-k})^2 = \\
 30 \quad &= \sum_n \left(\sum_{k=0}^M \bar{a}_k s_{n-k} \right)^2 \quad \text{met } \bar{a}_0 = 1
 \end{aligned}$$

Dit is te herleiden tot:

$$35 \quad E = \sum_{k=0}^M \bar{a}_k^2 r_k^2 + 2 \sum_{k=0}^M \bar{a}_k \sum_{j=k+1}^M \bar{a}_j r_{k-j}$$

waarin

$$5 \quad r_k = \sum_{j=1}^{n-k} s_j \cdot s_{j+k} ;$$

dit zijn de autokorrelatiecoëfficiënten die al berekend zijn en ook gediend hebben als input voor het Split Levinson algoritme.

10 In het minimaliseringsalgoritme wordt het minimum van de fout gezocht voor de bandbreedte van de eerste formant, vervolgens voor de tweede formant, enzovoorts en daarna weer opnieuw voor de eerste formant, enzovoorts. Dit proces wordt zolang herhaald tot de bandbreedte-
getallen niet meer veranderen. De waardes voor de bandbreedtes worden
15 genomen uit een tabel met een zekere kwantisatie. Deze kwantisatie is met verschillende stapgroottes getest zonder dat de convergentie ooit mislukte. De volgorde waarin de minimalisatie verloopt (hier voor achtereenvolgens formant 1, 2, 3, 4 en 5) is van belang voor de snelheid van convergentie.

20 Figuur 4 toont een stroomdiagram van de werkwijze volgens de uitvinding. De werkwijze wordt gestart in blok 40. In blok 41 wordt een in een zeker tijdinterval, van bij voorbeeld 25 ms lang, gelegen gedeelte van het spraaksignaal ingenomen. Het signaal wordt bewerkt onder invloed van een "Hamming"-venster. Vervolgens worden uit de dan
25 verkregen bemonsteringen S_1, \dots, S_N autokorrelatiecoëfficiënten r_i ($i=0, \dots, M$) waarbij $M \ll N$ berekend in het blok 42. In blok 43 wordt het Split Levinson algoritme toegepast, uitgaande van de autokorrelatiecoëfficiënten r_i . Na een aantal, te weten M , rekursiestappen in het Split Levinson algoritme heeft men de nulpunten
30 $np_{M,1}, np_{M,2}, \dots, np_{M, 1/2 M}$ (M is even). Vervolgens worden in het blok 44 uit de in de laatste rekursiestap verkregen nulpunten de formantfrequenties $f_1, \dots, f_{M/2}$ afgeleid. Daarna worden in blok 45 de bij de formantfrequenties behorende bandbreedtes B_1 tot en met $B_{M/2}$ afgeleid. Vervolgens loopt het programma via de keten 46, 47 weer
35 terug naar blok 41 en wordt uit een, over een zeker tijdinterval (van bij voorbeeld 10 ms) verschoven, tijdinterval (van 25 ms) een spraaksignaal afgeleid waaruit opnieuw een set formantfrequenties met

bijbehorende bandbreedtes kan worden afgeleid. Het programma wordt dus telkens herhaald totdat het volledige spraaksignaal is gekodeerd. Het programma eindigt dan via 46 en 48.

Figuur 5 toont een verdere uitwerking van blok 43 van
5 figuur 4. Figuur 5 toont een stroomschema van het Split Levinson algoritme zoals dat hiervoor al is uiteengezet. Het programma start in blok 50. In de blokken 51 en 52 wordt $P_0(z)$ respektievelijk $P_1(z)$ berekend. Het nulpunt van $P_1(z)$ $np_{1,1}$ ligt bij $z^{-1} = -1$. Vervolgens wordt $k = 2$ genomen (blok 53) en wordt in blok 54 de
10 singuliere prediktor polynoom $P_k(z)$ berekend volgens formule (8). Afhankelijk van de vraag of k even of oneven is (blok 55) worden de nulpunten $np_{k,1}$, $np_{k,2}$, òf volgens blok 56 of volgens blok 57 bepaald. Vervolgens wordt de waarde k met 1 verhoogd (blok 58) en loopt het programma via 59 en de keten 60 terug naar blok 54 voor het
15 doorlopen van de volgende rekursiestap. Na de laatste rekursiestap ($k = M$) loopt het programma via 59 naar blok 61 en is het programma beëindigd.

Figuur 6 toont een uitvoeringsvoorbeeld van de inrichting volgens de uitvinding voor het uitvoeren van de werkwijze. Via de
20 ingangsklem 65 wordt een spraaksignaal aan de inrichting aangeboden. In de eerste eenheid 66 wordt een in een zeker tijdinterval gelegen gedeelte van het spraaksignaal gebruikt voor het berekenen van een parameterwaarde, bij voorbeeld de autokorrelatiecoëfficiënt voor opeenvolgende tijdstippen liggend binnen dit tijdinterval. Deze
25 parameterwaardes worden toegevoerd aan een tweede eenheid 67. Deze eenheid 67 past het Split Levinson algoritme toe op de toegevoerde parameterwaardes. De in de laatste rekursiestap van het Split Levinson algoritme verkregen nulpunten worden toegevoerd aan de derde eenheid 68 die daaruit de formantfrequenties afleidt. Bovendien kan de derde
30 eenheid 68 zijn ingericht voor het berekenen van de bijbehorende bandbreedtes. De resultaten worden aan een uitgang 69 van de derde eenheid 68 aangeboden.

Het zij vermeld dat diverse modifikaties van de getoonde werkwijze respektievelijk inrichting mogelijk zijn, zonder dat wordt
35 afgeweken van datgene dat valt onder de scope van de uitvinding, zoals gedefinieerd in de conclusies.

Tabel -Inscripten in de stroomdiagrammen van figuren 4 en 5.

5	
bloknummer	inschrift
40, 50	start
41	derive speech segment
10 42	determine auto correlation coefficients
43	determination of zeroes in Split Levinson algorithm
44	calculate formant frequencies
45	calculate corresponding bandwidths
15 48	stop
54	calculate singular predictor polynomial
55	k even?
56, 57	determination of zeroes
61	return

LITERATUURLIJST

- 5
- (1) Philips' Elcoma technical publication no. 101 (1983)
MEA 8000 voice synthesizer: principles and interfacing.
- (2) Philips' Elcoma technical publication no. 217 (1986)
10 Speech synthesis: the complete approach with the PCF 8200.
- (3) Vogten, L.L.M. (1983) Analyse, zuinige kodering en resynthese
van spraakgeluid. Dissertatie, Eindhoven.
- 15 (4) Delsarte, P. and Genin, Y.V. (1986) The Split Levinson Algorithm.
IEEE Trans. on ASSP, Vol. ASSP-34, No. 3, June 86, p. 470-478.
- (5) Markel, J.D. and Gray, A.H. (1976) Linear prediction of speech
Springer Verlag.
- 20
- (6) Hildebrand, F.B., Introduction to numerical analysis.
McGraw Hill (1956).
- (7) Sugamura, N. en Itakura, F., Speech analysis and synthesis
25 methods developed at ELL in NTT - From LPC to LSP,
in Speech Communication Vol. 5, 1986, p. 199-215.

Conclusies:

1. Werkwijze voor het bepalen van formantfrequenties uit een binnen een zeker tijdsinterval liggend gedeelte van een spraaksignaal, waarbij
 - voor opeenvolgende tijdstippen liggend binnen het tijdsinterval uit het
 - 5 binnen het tijdsinterval liggend gedeelte van het spraaksignaal een parameterwaarde wordt afgeleid,
 - uit de parameterwaardes wordt een polynoom van een zekere orde bepaald,
 - uit de bepaalde polynoom worden de formantfrequenties afgeleid, met het kenmerk, dat een Split Levinson algoritme wordt uitgevoerd,
 - 10 waarbij in een aantal achtereenvolgende rekursiestappen uit de parameterwaardes telkens een singuliere prediktor polynoom van een hogere orde wordt bepaald, en dat na de laatste rekursiestap uit de in de laatste rekursiestap verkregen singuliere prediktorpolynoom de formantfrequenties worden afgeleid.
- 15 2. Werkwijze volgens conclusie 1, met het kenmerk, dat in een rekursiestap de nulpunten van de in deze rekursiestap bepaalde singuliere prediktorpolynoom worden afgeleid, daarbij gebruik makend van de nulpunten berekend tijdens de voorgaande rekursiestap, en dat na de laatste rekursiestap uit de in deze rekursiestap verkregen nulpunten de
- 20 formantfrequenties worden afgeleid.
3. Werkwijze volgens conclusie 1 of 2, met het kenmerk, dat bij elk der aldus gevonden formantfrequenties de bijbehorende bandbreedte wordt bepaald uitgaande van de parameterwaardes en de berekende formantfrequenties, door middel van een
- 25 minimaliseringsalgoritme.
4. Werkwijze volgens conclusie 1, 2 of 3, met het kenmerk, dat de parameterwaarde de waarde van de autokorrelatiekoëfficiënt is.
5. Inrichting voor het uitvoeren van de werkwijze volgens één der voorgaande conclusies, voorzien van
 - 30 - een ingangsklem voor het ontvangen van een spraaksignaal,
 - een eerste eenheid voor het, voor opeenvolgende tijdstippen liggend binnen een tijdsinterval, uit het binnen dit tijdsinterval gelegen gedeelte van het spraaksignaal afleiden van een parameterwaarde, met een ingang gekoppeld met de ingangsklem, en een uitgang,
 - 35 - een tweede eenheid voor het bepalen van een polynoom van een zekere orde uit de parameterwaardes, met een ingang gekoppeld met de uitgang van de eerste eenheid, en een uitgang, en

0005105

- een derde eenheid voor het afleiden van de formantfrequenties uit de bepaalde polynoom, met een ingang gekoppeld met de uitgang van de tweede eenheid en een uitgang voor het afgeven van de formantfrequenties, met het kenmerk, dat de tweede eenheid is ingericht voor het uitvoeren
5 van een Split Levinson algoritme, en voor het daarin in een aantal achtereenvolgende rekursiestappen uit de parameterwaardes afleiden van telkens een singuliere prediktor polynoom van een hogere orde en dat de derde eenheid is ingericht voor het afleiden van de formantfrequenties uit de in de laatste rekursiestap verkregen singuliere prediktor
10 polynoom.

6. Inrichting volgens conclusie 5, voor het uitvoeren van de werkwijze volgens conclusie 2, met het kenmerk, dat de tweede eenheid verder is ingericht voor het in een rekursiestap afleiden van de nulpunten van de in deze rekursiestap bepaalde singuliere
15 prediktor polynoom, daarbij gebruik makend van de nulpunten berekend tijdens de voorgaande rekursiestap, en dat de derde eenheid ingericht is voor het afleiden van de formantfrequenties uit de nulpunten verkregen in de laatste rekursiestap.

7. Inrichting volgens conclusie 5, voor het uitvoeren van de
20 werkwijze volgens conclusie 3, met het kenmerk, dat de derde eenheid verder is ingericht voor het bij elk der aldus gevonden formantfrequenties bepalen van de bijbehorende bandbreedte, uitgaande van de parameterwaardes en de berekende formantfrequenties, door middel van een minimaliseringsalgoritme.

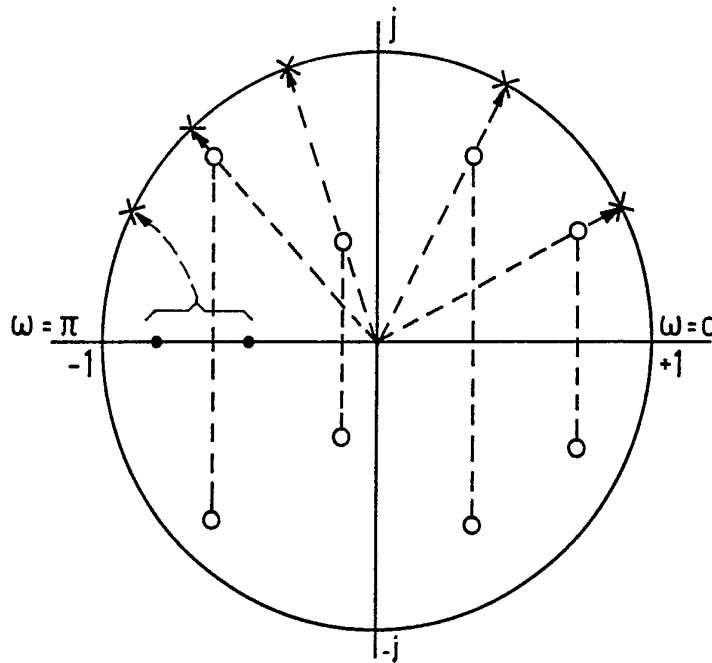


FIG. 1

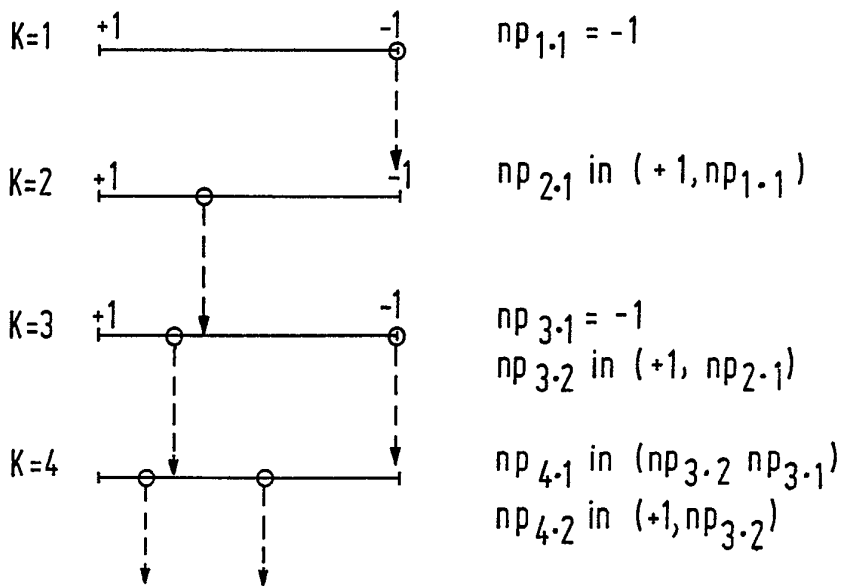


FIG. 2

enz.

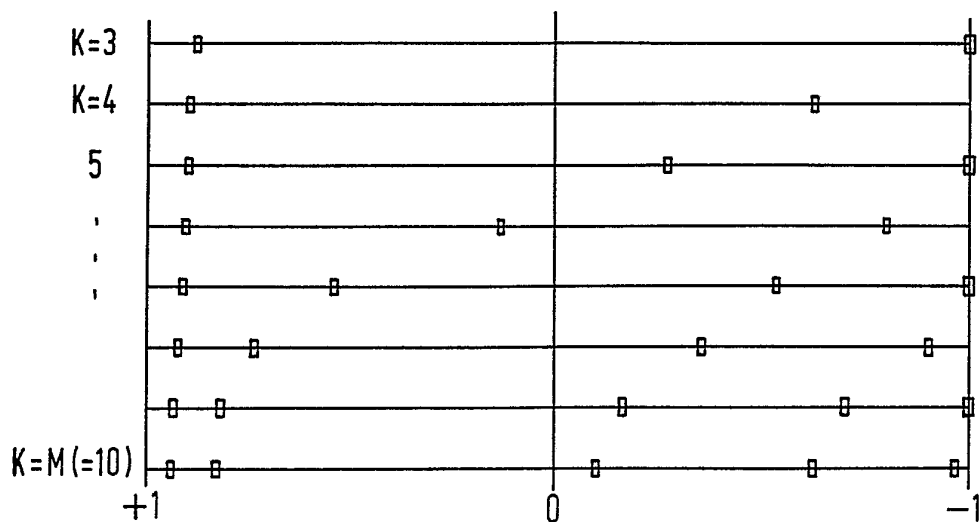


FIG. 3

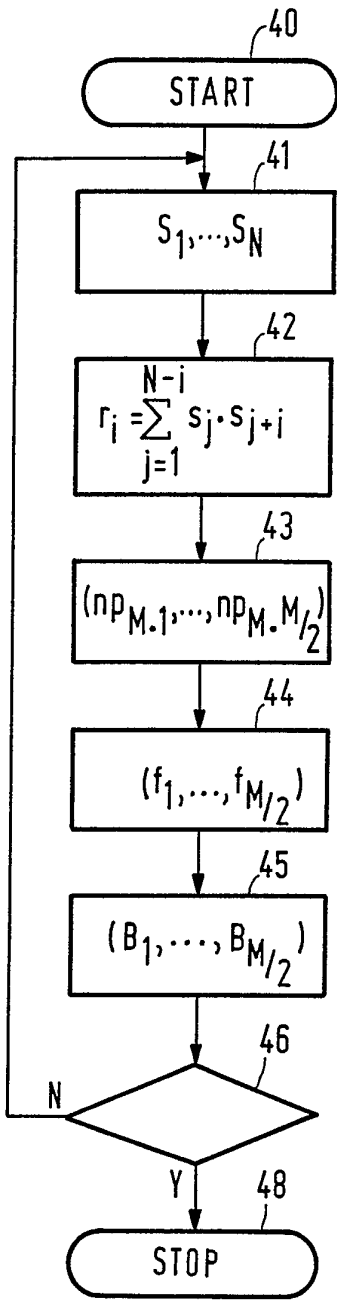


FIG. 4

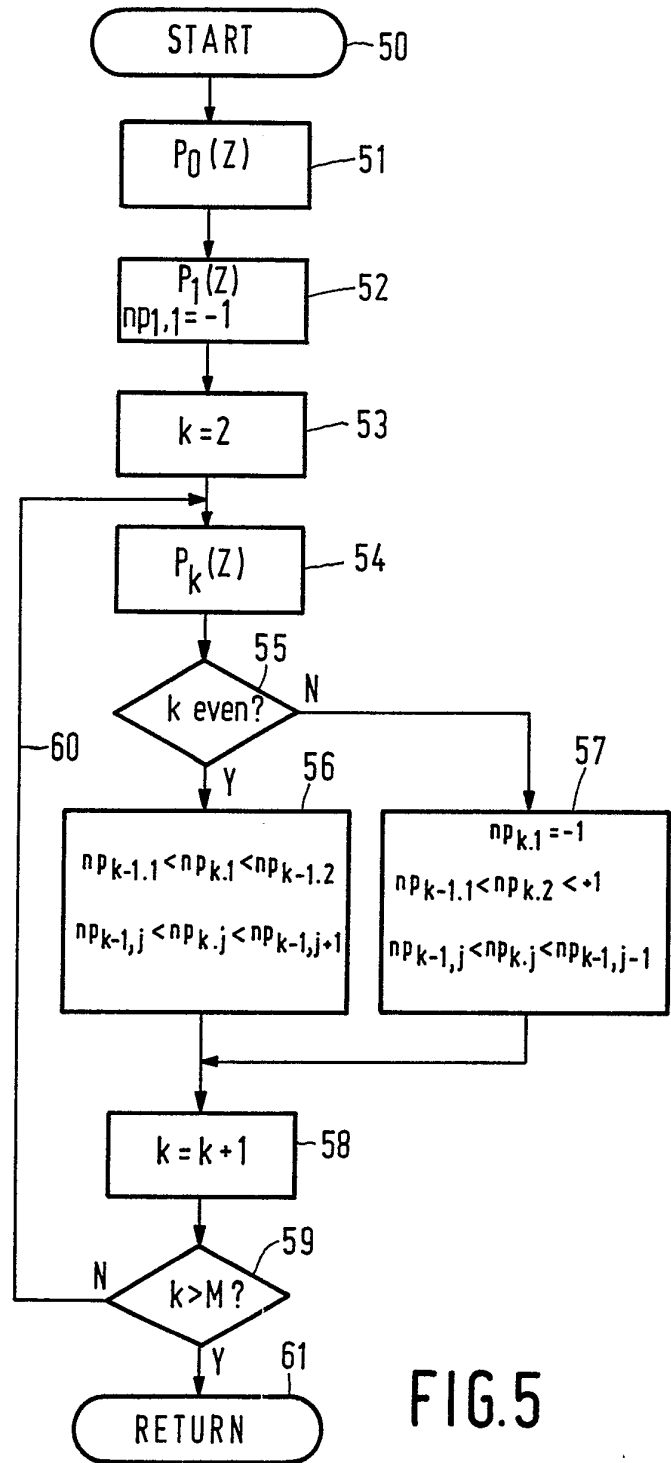


FIG. 5

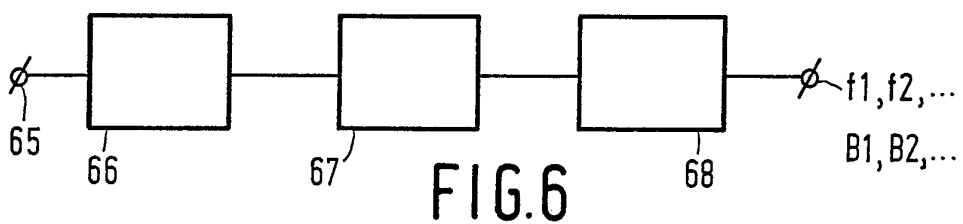


FIG. 6