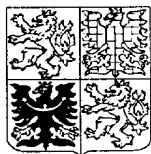


PATENTOVÝ SPIS

(11) Číslo dokumentu:

286 743

(19)
ČESKÁ
REPUBLIKA



ÚŘAD
PRŮMYSLOVÉHO
VLASTNICTVÍ

(21) Číslo přihlášky: 1996 - 678

(22) Přihlášeno: 14.09.1994

(30) Právo přednosti:

14.09.1993 EP 1993/93307211

06.12.1993 GB 1993/9324967

21.06.1994 GB 1994/9412451

(40) Zveřejněno: 17.07.1996

(Věstník č. 7/1996)

(47) Uděleno: 20.04.2000

(24) Oznámeno udělení ve Věstníku: 14.06.2000

(Věstník č. 6/2000)

(86) PCT číslo: PCT/GB94/01999

(87) PCT číslo zveřejnění: WO 95/08170

(13) Druh dokumentu: B6

(51) Int. Cl.⁷:

G 10 L 11/00

H 04 Q 1/46

(73) Majitel patentu:

BRITISH
TELECOMMUNICATIONS PUBLIC LIMITED
COMPANY, London, GB;

(72) Původce vynálezu:

Barrett Paul Alexander, Suffolk, GB;

(74) Zástupce:

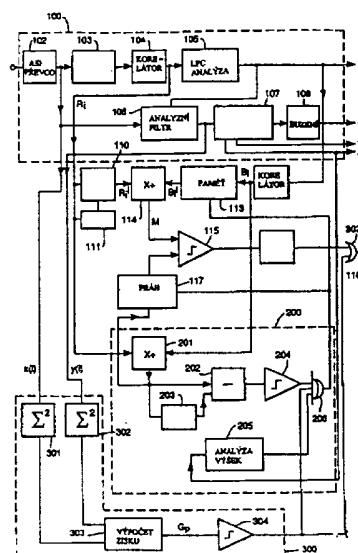
Guttman Michal JUDr. Ing., Nad Štolou 12,
Praha 7, 170 00;

(54) Název vynálezu:

Detektor hlasu

(57) Anotace:

Detektor hlasu pro zjišťování přítomnosti řeči ve vstupním signálu, který má šumovou složku, sestávající z prostředků (113) pro uchování modelu šumové složky vstupního signálu, z prostředků (114, 115) pro zjištění spektrální podobnosti vstupního signálu s uloženým modelem šumu pro vytváření výstupního rozhodovacího signálu, z prostředků (112) pro aktualizaci uloženého modelu šumu, a z pomocného detektoru (200) pro řízení prostředků (112) pro aktualizaci tak, aby k aktualizaci došlo pouze v době, kdy pomocný detektor (200) udává, že vstupní signál neobsahuje řeč. Detektor hlasu dále zahrnuje prostředky (303) pro výpočet parametru předurčeného zisku vstupního signálu a modifikační prostředky (304) pro potlačení aktualizace v případě, že předurčený zisk překročí prahovou hodnotu.



CZ 286743 B6

Detektor hlasu

Oblast techniky

5

Vynález se týká detektoru hlasu pro zjišťování přítomnosti řeči ve vstupním signálu, který má šumovou složku, sestávající z prostředků pro uchování modelu šumové složky vstupního signálu, z prostředků pro zjištění spektrální podobnosti vstupního signálu s uloženým modelem šumu pro vytváření výstupního rozhodovacího signálu, z prostředků pro aktualizaci uloženého modelu šumu, a z pomocného detektoru pro řízení prostředků pro aktualizaci tak, aby k aktualizaci došlo pouze v době, kdy pomocný detektor udává, že vstupní signál neobsahuje řeč.

10

Dosavadní stav techniky

15

Detektor hlasu je zařízení, do něhož je přiváděn vstupní signál s cílem stanovit úseky tohoto signálu obsahující řeč nebo úseky obsahující pouze šum. Jedním zvláště významným použitím takových detektorů hlasu, na něž však tento vynález není omezen, jsou mobilní radiové telefonní systémy, kde údaje o přítomnosti či nepřítomnosti řeči v signálu mohou být využívány pro snížení spotřeby energie a pro snížení interference způsobené vypínáním telefonu v úsecích ticha. U těchto mobilních systémů je v důsledku umístění jednotek ve vozidlech obvykle vysoká úroveň šumu. Další možností je použití u radiových systémů pro zlepšení využití radiového spektra.

20

Na obr. 1 je zobrazen známý detektor hlasu tak, jak je popsán v PCT patentové přihlášce č. WO89/08910.

25

Na jeho vstupu 1 jsou přijímány signály. Paměť 2 pak obsahuje údaje definující předpoklad nebo model frekvenčního spektra šumu. Tyto údaje jsou následně porovnávány v komparátoru 3 se spektrem přicházejícího signálu a získá se míra jejich podobnosti, která se dále porovnává v prvku 4 s prahovou hodnotou. Pro sledování změn v šumové složce je model šumu aktualizován ze vstupu, a to pouze v době nepřítomnosti řeči ve vstupním signálu. Rovněž uvedená prahová hodnota může být upravována adaptérem 6.

30

S cílem zaručit, že k úpravám modelu šumu dojde jenom v obdobích, kdy vstupní signál obsahuje pouze šum, bez nebezpečí postupné nesprávné aktualizace, která by následovala po špatném rozhodnutí, je jeho úprava prováděna za řízení pomocným detektorem 7, který zahrnuje detektor 8 nehlasových zvuků a detektor hlasových zvuků 9. Pomocný detektor 7 považuje řeč za přítomnou ve vstupním signálu, jestliže oba detektory 8 a 9 zjistí zvuk. Po tomto zjištění je následně umožněna aktualizace šumového modelu i úprava prahové hodnoty hlavního detektoru. Detektor 8 nehlasových zvuků získává sadu LPC koeficientů signálu a porovnává autokorelační funkci těchto koeficientů mezi po sobě následujícími snímky, zatímco detektor 9 hlasových zvuků zkoumá změny autokorelace LPC zbytkového signálu.

40

Toto uspořádání je velmi úspěšné pro rozlišování úseků řeči a úseků v nichž je přijímán pouze šum. Problém je však v tom, že pomocný detektor 7 považuje za běžný šum i signální tóny (tj. pomocný detektor 7 v nich nerozpozná řeč) tak, že hlavní detektor aktualizuje model šumu podle těchto tónů, jako kdyby byly šumem, a dále v tom, že dojde k zabránění přenosu signálních tónů, nebo alespoň k předčasnému přerušení jejich přenosu.

45

Tento problém by bylo možno překonat přidáním tónových detektorů, přičemž každý by byl naladěný na frekvenci konkrétního signálního tónu. Rozmanitost jednotlivých signálních tónů ve světě je však značná, takže by bylo potřeba velkého množství jednotlivých detektorů, aby

50

například uživatel mobilního telefonu uskutečňující mezinárodní hovor mohl slyšet obsazovací tón nezávisle na zemi v které vznikl.

5 Podstata vynálezu

Uvedené nedostatky odstraňuje detektor hlasu pro zjišťování přítomnosti řeči ve vstupním signálu, který má šumovou složku, sestávající z prostředků pro uchování modelu šumové složky vstupního signálu, z prostředků pro zjištění spektrální podobnosti vstupního signálu s uloženým modelem šumu pro vytváření výstupního rozhodovacího signálu, z prostředků pro aktualizaci uloženého modelu šumu, a z pomocného detektoru pro řízení prostředků pro aktualizaci tak, aby k aktualizaci došlo pouze v době, kdy pomocný detektor udává, že vstupní signál neobsahuje řeč, podle vynálezu, jehož podstata spočívá v tom, že dále zahrnuje prostředky pro výpočet parametru předurčeného zisku vstupního signálu a modifikační prostředky pro potlačení aktualizace v případě, že předurčený zisk překročí prahovou hodnotu.

Podle výhodného provedení pomocný detektor zahrnuje detektor hlasových zvuků reagující na signály odvozené z kódování s lineární predikcí zbytkového signálu.

20 Podle dalších výhodných provedení jsou prostředky pro výpočet parametru předurčeného zisku upraveny pro výpočet předurčeného zisku kódování s lineární predikcí předurčení šestého nebo nižšího řádu, případně čtvrtého nebo nižšího řádu.

25 Podle dalšího výhodného provedení detektor hlasu dále zahrnuje prostředky pro zjišťování jedné nebo více primárních frekvenčních složek vstupního signálu a pro porovnání frekvence s předem stanovenou prahovou hodnotou a pro umožnění potlačení aktualizace pouze v případě, že uvedená primární složka překračuje prahovou hodnotu.

30 Před prostředky pro výpočet parametru předurčeného zisku může být předřazen filtr pro odstranění horní části frekvenčního rozsahu vstupního signálu, přičemž prostředky pro výpočet parametru předurčeného zisku jsou upraveny pro výpočet parametru předurčeného zisku pouze pro frekvenční složky ležící v pásmu propustnosti filtru.

35 Podle jiného výhodného provedení má detektor hlasu dva filtry s příslušnými pásmy propustnosti a příslušné prostředky pro výpočet parametrů předurčeného zisku u příslušných pásmech propustnosti a modifikační prostředky pro potlačení aktualizace v případě, že předurčený zisk přesahuje v obou pásmech propustnosti prahovou hodnotu.

40 Je též výhodné, pokud detektor hlasu zahrnuje prostředky pro vzorkování filtrovaného signálu/signálů.

45 Podstata řešení podle tohoto vynálezu pak spočívá v tom, že detektor hlasu zahrnuje prostředky pro výpočet parametru předurčeného zisku vstupního signálu a dále modifikační prostředky uspořádané pro znemožnění aktualizace uloženého modelu šumu v případě, že předurčený zisk překročí prahovou hodnotu.

Objasnění obrázků na výkresech

50 Následně budou popsána některá provedení tohoto vynálezu s odkazem na přiložené výkresy, kde:

obr. 1 zobrazuje blokové schéma známého detektoru hlasu podle PCT přihlášky č. WO89/08910,

- obr. 2 představuje blokové schéma zvukového kodéru s detektorem hlasu podle tohoto vynálezu,
 obr. 3 a obr. 4 ukazují graficky hodnoty předurčeného zisku u různých vstupních signálů,
 5 obr. 5, obr. 6 a obr. 7 představují bloková schémata dalších provedení tohoto vynálezu.

Příklady provedení vynálezu

- 10 Běžný hlasový kodér 100 na obr. 2 má zvukový vstup 101. Zvukový signál je zde vzorkován při 8 kHz a převáděn do digitální formy analogově-digitálním převodníkem 102. Jednotka 103 rozděluje zvukové vzorky do snímků, například se 160 vzorky (tj. do 20ms vzorků) a multipluje je funkcí, která snižuje podíl vzorků na začátku a na konci snímku. Korelátor 104 přijímá digitalizované zvukové vzorky a vytváří pro každý snímek autokorelační koeficienty R_i . Jednotka
 15 105 LPC analýzy (LPC=Linear Predictive Coding=kódování s lineární predikcí) vypočítává za použití známé metody, např. Levinson-Durbinovým nebo Schurrovým algoritmem, koeficienty a_i filtru (někdy označovaného jako syntézní filtr), majícího frekvenční charakteristiku, která odpovídá frekvenčnímu spektru vstupního zvukového signálu.
- 20 Digitalizovaný vstupní signál prochází rovněž inverzním filtrem, nebo analyzním filtrem 106, ovládaným koeficienty, pro vytvoření zbytkového signálu, který je dále analyzován jednotkou 107 dlouhodobého předurčení, která vypočítává optimální zpoždění pro předurčení LPC zbytkového signálu z jeho předchozích hodnot a odpovídající hodnotu zisku pro předurčení. Analyzní filtr 106 vytváří také druhý zbytkový signál (tj. rozdíl mezi současným LPC zbytkovým
 25 signálem a LPC zbytkovým signálem po zpoždění a vyvážením získanými parametry). Budící jednotka 108 získává parametry pro přenos do dekodéru ze zbytkového signálu dlouhodobého předurčení (LTP), nebo jinými známými prostředky.

LPC koeficienty a_i , zpoždění d , zisk g a budící parametry e jsou přenášeny do dekodéru.

- 30 Hlavní detektor hlasu podle dříve zmíněné patentové přihlášky zjišťuje průměr autokorelačních koeficientů R_i pomocí nivelizačního zařízení 110, které vytváří vážený součet R_i' současných koeficientů R_i a koeficientů z předchozích snímků, uložených v paměti 111. Prostředky 112 pro aktualizaci uloženého modelu šumu, tvořené dalším autokorelátorem, vytváří autokorelační
 35 koeficienty B_i těchto LPC koeficientů a_i , které přecházejí do paměti, která tvoří prostředky 113 pro uchovávání modelu šumové složky. Obsah této paměti je aktualizován pouze v průběhu úseků vstupního signálu, o nichž pomocný detektor, který bude popsán níže, předpokládá, že obsahují pouze šum. Obsah paměti B_i pak představuje odhad šumového spektra vstupního signálu. Násobná/součtová jednotka 114, která je součástí prostředků pro zjištění spektrální
 40 podobnosti vstupního signálu s uloženým modelem šumu, pak vytváří míru M spektrální podobnosti mezi vstupním signálem a šumovým modelem, která je definována jako

$$M = B_0' + 2 \sum_{i=1}^n \frac{R_i' B_i'}{R_0'}$$

- 45 Nulový spodní index značí počáteční autokorelační koeficient a n je počet vzorků ve zvukovém snímku.

- Míra M spektrální podobnosti se pak v komparátoru 115, který je součástí prostředků pro zjištění spektrální podobnosti vstupního signálu s uloženým modelem šumu, porovnává s prahovou
 50 hodnotou a na výstupu 116 se vytváří signál indikující přítomnost nebo nepřítomnost řeči ve

vstupním signálu. Prahová hodnota může být přizpůsobitelně nastavena prvkem 117 podle aktuální úrovně šumu.

5 Aktualizace modelu šumu uloženého v paměti prostředků 113 pro uchovávání modelu šumové složky není ovládána výstupem 116 právě popsaného detektoru, neboť chyba v rozpoznání řeči by měla za následek aktualizaci paměti informací o řeči a následně další chyby v rozpoznávání. Z tohoto důvodu je aktualizace řízena pomocným detektorem 200. Aby bylo možno odlišit šum a nehlasový zvuk je vytvářen v bloku 201 součet (nezprůměrovaných) autokorelačních koeficientů B_i LPC koeficientů. Odčítač 202 porovnává tento součet s odpovídajícím součtem u předcházejícího zvukového snímku uloženého v paměti 203. Tento rozdíl, představující spektrální podobnost mezi po sobě následujícími snímky vstupního signálu, je pro vytvoření rozhodovacího signálu porovnán s prahovou hodnotou v porovnávací jednotce 204.

15 Pro rozpoznání hlasového zvuku se detektorem 205 hlasových zvuků měří zpoždění d . Výstup detektoru 205 hlasových zvuků je v součtovém hradle 206 sloučen s výstupem porovnávací jednotky 204. Pomocný detektor 200 indikuje přítomnost hlasu ve vstupním signálu, jestliže každá z jednotek 204 a 205 vytváří výstupní signál indikující přítomnost hlasu. Jak bylo uvedeno již v úvodu, má-li systém propouštět signální tóny, musí tyto rozpoznávat spíše jako řeč než jako šum. Právě popsaný pomocný detektor však není pro dosažení tohoto cíle příliš efektivní. 20 Ačkoliv rozpoznává určité signální tóny, jiné (obecně ty, které mají relativně chudý spektrální obsah) nejsou rozpoznány. Když dojde k chybě pomocného detektoru 200, dojde rovněž k chybě hlavního detektoru, neboť model šumu v paměti je pak adaptován podle signálního tónu.

25 Pro zjišťování signálních tónů ve vstupním signálu je proto vytvořen další pomocný detektor. Výhodně je využíváno zjištění, že signální tóny, které jsou uměle generovány, obsahují malý počet kmitočtových složek (které mohou být modulovány). Pro takové signály je použitím rozlišení mezi tónově založenými signály (včetně vícetónových signálů) a pozadím nebo šumovými signály okolí výkon LPC předurčovače zvláště vysoký.

30 LPC předurčený zisk G_p je definován jako poměr výkonu vstupního signálu a výkonu výstupního signálu u hlasového snímku

$$G_p = \frac{\sum_{i=0}^{n-1} x^2(i)}{\sum_{i=0}^{n-1} y^2(i)},$$

35 kde x_i je vstup filtru a y_i je výstup inverzního filtru:

$$y(t) = x(t) + \sum_{i=1}^m y(t-i) a_i$$

40 kde m je počet koeficientů filtru, který bývá obvykle 8 nebo 10. Signály $x(i)$ a $y(i)$ se získávají z LPC hlasového kodéru 100, na výstupech analogově-digitálního převodníku 102 resp. analýzního filtru 106. Tyto hodnoty jsou umocněny v umocňovacích jednotkách 301 a 302 a předurčený zisk se stanoví v prostředcích 303 pro výpočet parametru předurčeného zisku vstupního signálu, které jsou tvořeny aritmetickou jednotkou, která vypočítává G_p podle shora uvedené rovnice. Výstup této jednotky je v komparátoru, který tvoří modifikační prostředky 304,

porovnáán s pevnou prahovou hodnotou T . Jestliže předurčený zisk přesahuje tuto prahovou hodnotu (obvykle $T = 63$ dB nebo $T = 18$ dB), je tón pokládán za rozpoznáný. Existuje několik možných reakcí na rozpoznání tónu:

- 5 a) nahradit výstup hlavního detektoru prostředky 303 pro výpočet parametru předurčeného zisku vstupního signálu,
- b) nahradit pomocný detektor prostředky třetího vstupu do součtového hradla 206,
- 10 c) obě předešlé možnosti.

15 Samozřejmě místo výpočtu podílu může být výraz Σx^2 porovnáván s výrazem Σy^2 , vynásobeným prahovou hodnotou. Na obr. 3 jsou zobrazeny histogramy předurčených zisků v dB, které byly získány ze šumu pozadí, z řeči, ze šumu pozadí v signálních tónech a ze samotných signálních tónů. Na obr. 4 jsou pak zobrazeny diagramy předurčeného zisku v závislosti na čase u různých signálních tónů používaných ve Velké Británii, konkrétně u:

- obsazovacího tónu,
- oznamovacího tónu,
- 20 vyzváněcího tónu,
- tónu "volané číslo neexistuje", a
- tónu "zařízení obsazeno".

25 V praxi jsou obsazovací tón, oznamovací tón a tón "volané číslo neexistuje" úspěšně rozpoznávány dalším detektorem, stejně jako multifrekvenční tóny. Vyzváněcí tón a tón "zařízení obsazeno" jsou rozpoznávány detektorem 205 hlasových zvuků.

30 Další detektor 300 lze považovat za detektor určitých druhů tónů, alternativně (u provedení z obr. 2) může být považován za detektor situace, kdy zbytkový signál y_i je tak malý, že funkce jednotky 107 dlouhodobého předurčení (a tím detektoru 205 hlasových zvuků) není silná.

35 Alternativní možností zjišťování hlasového zvuku je nahradit detektor 205 hlasových zvuků jednotkami analogickými k jednotkám 301, 302, 303 a 304 a založit vytváření předurčeného zisku na analýze dlouhodobého předurčení v jednotce 107 dlouhodobého předurčení.

40 Dále budou s odkazem na obr. 5 popsány dvě modifikace zařízení z obr. 2. Vypočítávaný zisk je u provedení z obr. 2 ziskem LPC analýzy hlasového kodéru 100, která by mohla obvykle používat předurčovač osmého nebo desátého řádu. Při vědomí toho, že základem této části analýzy je, že informační tóny rezultují ve větších předurčených ziscích než šum prostředí a že čím vyšší je řád analýzy, tím vyšší je schopnost předurčovače modelovat šum prostředí, bylo zjištěno, že omezením výpočtu zisku na analýzu čtvrtého řádu dávají informační signály, sestávající z jednoho nebo dvou tónů, vysoký předurčený zisk, zatímco předurčený zisk u šumu prostředí může být snížen.

45 Principiálně by toto bylo možno dosáhnout použitím analýzy čtvrtého řádu a použitím filtru vedle jednotek 105, 106 pro zásobování pomocného detektoru signálem, avšak snazší je vypočítat předurčený zisk z reflexních koeficientů (někdy uváděných jako Parcor koeficienty). Na obr. 5 jsou tyto reflexní koeficienty vypočítávány známým způsobem z autokorelačních koeficientů R_i kalkulátorem 400 reflexních koeficientů (ačkoliv v závislosti na provedení 50 zvukového kodéru by mohlo být možno snímat je z mezilehlého bodu jednotky 105 LPC analýzy). Míru předurčeného zisku, chybu předurčení P_e , lze získat výpočtem z prvních čtyř reflexních koeficientů R_c podle vzorce:

$$Pe = \prod_{i=1}^4 (1 - Rc_i^2)$$

5 Tento výpočet se provádí v jednotce 401. Velká chyba předurčení Pe odpovídá nízkému předurčenému zisku a naopak. Signální tón se proto považuje za přítomný ve vstupním signálu, jestliže hodnota Pe je menší než prahová hodnota P_{th} . Toto srovnání v komparátoru 403 nahradilo srovnávání v modifikačních prostředcích 304 z obr. 2.

10 Šum okolního prostředí mobilního telefonu obsahuje velmi silné rezonance na nízkých frekvencích. Pro zjištění, zda je "tón" pod prahovou frekvencí se provádí další test. Volba prahové hodnoty v sobě zahrnuje určitý stupeň kompromisu, avšak doporučována je hodnota 385 Hz. neboť většina signálních tónů leží pod 400 Hz.

15 Tento další test pracuje tak, že jsou stanoveny pólové frekvence LPC filtru. Pro zmenšení komplexnosti analýzy je výhodný filtr nízkého řádu. Následně by mohla být provedena další LPC analýza, ale jednodušší je pokračovat, jako u provedení na obr. 5, výpočtem LPC koeficientů z reflexních koeficientů. Za předpokladu, že se použijí pouze první dva reflexní koeficienty z kalkulátoru 400 reflexních koeficientů, jsou LPC koeficienty a_i vypočítány běžným způsobem prvkem 404, který je definován tak, aby citlivost syntézního filtru byla:

$$20 \quad H(z) = 1 / (a_0 + a_1 z^{-1} + a_2 z^{-2})$$

Pak jsou polohy pólů v rovině z dány řešením kvadratické rovnice

$$a_0 z^2 + a_1 z + a_2 = 0 \quad a_0 = 1$$

$$z = \frac{-a_1}{2} + j \sqrt{\frac{4a_2 - a_1^2}{4}}$$

25 Jestliže výraz pod druhou odmocninou je záporný, pak póly leží na reálné ose a vstupní signál není signálním tónem. Jestliže je tento výraz kladný, ale reálná část polohy pólu je záporná (tj. $a_1 < 0$), pak je pól v levé polovině roviny z . To nezbytně znamená, že frekvence je více než 25 % hodnoty vzorkovací frekvence, tj. více než 2000 Hz pro vzorkovací frekvenci $f_s = 8$ kHz. V takovém případě je výpočet zbytečný a signál "> 385" může být ihned generován.

Pólová frekvence je dána rovnicí:

$$f = \arctan \left(\frac{-\sqrt{4a_2 - a_1^2}}{a_1} \right) \times \frac{f_s}{2\pi}$$

35 Podmínku, že $f < 385$ Hz lze zapsat (vyloučením druhých odmocnin jako:

$$(4a_2 - a_1^2) / a_1^2 < \tan^2 \left(\frac{2\pi \times 385}{f_s} \right)$$

nebo jako:

$$(4a_2 - a_1^2) / a_1^2 < 0,0973 \quad \text{při } f_s = 8\text{kHz}$$

- 5 Tento výpočet se provádí v jednotce prostředků 405 pro zjišťování jedné nebo více primárních frekvenčních složek vstupního signálu. Její výstup je sloučen v AND hradle 406 s výstupem komparátoru 403, takže rozhodnutí o tom, že se jedná o signální tón vznikne pouze, když předurčený zisk je velký a zároveň když je pólová frekvence větší než 385 Hz.
- 10 Jestliže je to požadováno, může být pólová frekvence vyšší než 2000 Hz (nebo vyšší než nějaký jiný horní limit) také odloučena, aby vysoké frekvence nad předpokládaným rozsahem signálních tónů nemohly být rozpoznány jako signální tóny.

- 15 Jestliže se při řešení kvadratické rovnice použije zvláštního výpočtu, je možné použít také třetí a čtvrtý reflexní koeficient. V tomto případě by mohly být zjištěny dva sdružené komplexní páry pólů s dvěma frekvencemi. Je zřejmé, že v takovém případě by signální tón nebyl pokládán za přítomný ve vstupním signálu, kdyby obě frekvence byly pod prahovou hodnotou.

- 20 Bylo již uvedeno, že provedení vynálezu podle obr. 2 a obr. 5 používají před autokorelačním výpočtem v jednotce 103 rozdělení zvukových vzorků do snímků (jak je běžné u LPC analýzy založené na autokorelaci). Jestliže se u hlasového kodéru nepožaduje provádět tuto činnost, je možnou alternativou případu z obr. 5 vypustit jednotku 103 a nahradit výpočet reflexního koeficientu konverzí autokorelačních hodnot na kovariantní hodnoty, jednotky 401, 404 modifikovat pro využití těchto kovariantních hodnot oproti využití reflexních koeficientů.
- 25 Alternativně, jak ukazuje obr. 6 (kde jsou zobrazeny pouze ty části, které byly oproti obr. 5 změněny), může být počáteční zpracování provedeno jednotkou 109 kovariantní analýzy, jejichž výstup je přiváděn do kalkulátoru 400' reflexního koeficientu a do autokorelační jednotky 104' modifikace autokorelačního koeficientu. Jednotka 105 LPC analýzy může být spojena, jako předtím s autokorelační jednotkou 104', nebo, jak je zobrazeno, přímo s jednotkou 109
- 30 kovariantní analýzy.

- Výše popsaná provedení pro detekci signálních tónů mají dobré výsledky, avšak mohou selhat u mechanicky generovaných tónů, které se používají v určitých teritoriích. Tyto mechanicky generované tóny mají vyšší harmonickou složku rezultující v malém předurčeném zisku. Prosté
- 35 odfiltrování vyšších harmonických není řešením, neboť vložení filtru vede ke zvýšení autokorelace všech signálů a tím k větším předurčeným ziskům rovněž u ostatních signálů. Bylo zjištěno, že předurčovač má snahu spíše upravovat póly filtru než charakteristiky vstupního signálu. Nicméně bylo také zjištěno, že dobré výsledky je možno získat filtrací, jestliže analýza předurčeného zisku je omezena na stanovení předurčení signálu pouze ve frekvenčním rozsahu
- 40 odpovídajícím pásmu propustnosti harmonického filtru. Toho se dosáhne vzorkováním signálu frekvencí dvojnásobnou oproti šířce pásma filtru před analýzou předurčeného zisku.

- Provedení vynálezu zobrazené na obr. 7, které je podobné provedení z obr. 5, proto používá filtr 450, kterým je nízkopropustný FIR filtr, mající pásmo propustnosti do 600 (3 dB) a mající zeslabení potlačovaného pásma 20 dB při 1200 Hz. Výhodné je, když zeslabení potlačovaného pásma není příliš velké. Výstup filtru 450 je vzorkován při 1200 Hz ve vzorkovací jednotce 451.

- U tohoto filtrování jsou velmi redukovány možnosti tónové detekce sdílet komponenty s hlasovým kodérem 100. Do filtru 450 je tudíž digitalizovaný vstupní signál přiváděn přímo
- 50 z analogově-digitálního převodníku 102 a jeho výstup směřuje do kalkulátoru 400' reflexních koeficientů nebo do jednotky kovariantní nebo autokorelační analýzy tak, jak bylo popsáno dříve.

Další provedení vynálezu zmírňuje problém harmonických složek bez nadměrného omezování frekvenčního pásma analýzy předurčovaného zisku. Toho je dosaženo použitím filtrů rozdělujících vstupní signál na dvě nebo více frekvenčních pásem, kde každé toto pásmo je dostatečně úzké na to, aby nemohlo obsahovat základní a třetí harmonickou signálního tónu. Každý kanál je pak vzorkován a podroben oddělené analýze předurčeného zisku.

U provedení podle obr. 8 je signál rozdělen filtry 450a a 450b na dvě frekvenční pásma 400 až 1200 Hz a 1200 až 2000 Hz a vzorkován při 1,6 kHz v prostředcích 451a, 451b pro vzorkování. Výpočet reflexního koeficientu 400'a a 400'b, analýza chyby předurčení 401a, 401b a prahování jsou prováděny odděleně pro obě pásma. Dva výstupy z komparátorů, které tvoří prostředky 403a a 403b pro potlačení aktualizace, jsou vedeny na oddělené vstupy součtového hradla 206 tak, aby vysoký předurčený zisk v obou kanálech byl pokládán za indikaci přítomnosti signálního tónu. Ostatní prvky 100 až 303 z obr. 7 nejsou u na obr. 8 zobrazeny, neboť jsou u obou provedení stejné.

PATENTOVÉ NÁROKY

1. Detektor hlasu pro zjišťování přítomnosti řeči ve vstupním signálu, který má šumovou složku, sestávající z prostředků (113) pro uchování modelu šumové složky vstupního signálu, z prostředků (114, 115) pro zjištění spektrální podobnosti vstupního signálu s uloženým modelem šumu pro vytváření výstupního rozhodovacího signálu, z prostředků (112) pro aktualizaci uloženého modelu šumu, a z pomocného detektoru (200) pro řízení prostředků (112) pro aktualizaci tak, aby k aktualizaci došlo pouze v době, kdy pomocný detektor (200) udává, že vstupní signál neobsahuje řeč, **v y z n a ě u j í c í s e t í m**, že dále zahrnuje prostředky (303) pro výpočet parametru předurčeného zisku vstupního signálu a modifikační prostředky (304) pro potlačení aktualizace v případě, že předurčený zisk překročí prahovou hodnotu.
2. Detektor hlasu podle nároku 1, **v y z n a ě u j í c í s e t í m**, že pomocný detektor (200) zahrnuje detektor (205) hlasových zvuků reagující na signály odvozené z kódování s lineární predikcí zbytkového signálu.
3. Detektor hlasu podle nároku 1 nebo 2, **v y z n a ě u j í c í s e t í m**, že prostředky (303) pro výpočet parametru předurčeného zisku jsou upraveny pro výpočet předurčeného zisku kódování s lineární predikcí předurčení šestého nebo nižšího řádu.
4. Detektor hlasu podle nároku 3, **v y z n a ě u j í c í s e t í m**, že prostředky (303) pro výpočet parametru předurčeného zisku jsou upraveny pro výpočet předurčeného zisku kódování s lineární predikcí předurčení čtvrtého nebo nižšího řádu.
5. Detektor hlasu podle kteréhokoliv z předchozích nároků, **v y z n a ě u j í c í s e t í m**, že dále zahrnuje prostředky (405) pro zjišťování jedné nebo více primárních frekvenčních složek vstupního signálu a pro porovnání frekvence s předem stanovenou prahovou hodnotou a pro umožnění potlačení aktualizace pouze v případě, že uvedená primární složka překračuje prahovou hodnotu.
6. Detektor hlasu podle kteréhokoliv z předchozích nároků, **v y z n a ě u j í c í s e t í m**, že před prostředky (303) pro výpočet parametru předurčeného zisku je předřazen filtr (450) pro odstranění horní části frekvenčního rozsahu vstupního signálu, přičemž prostředky (303) pro

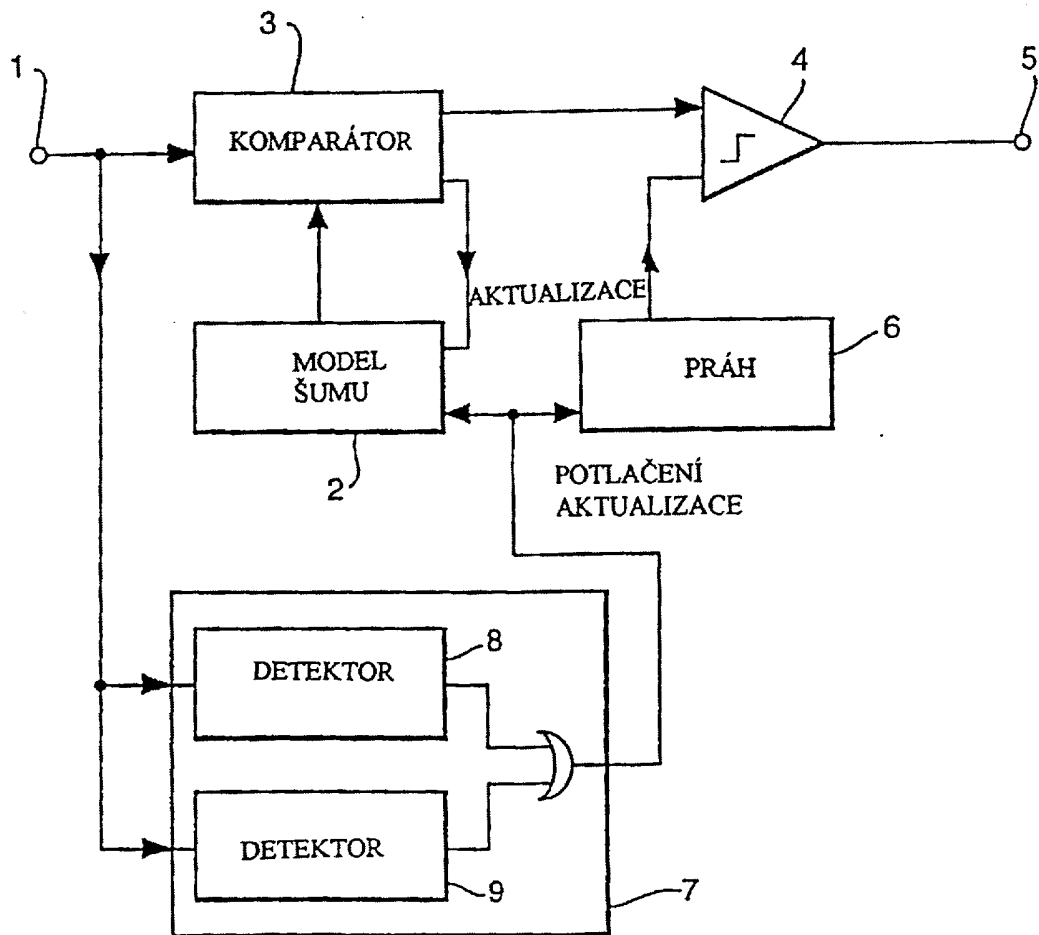
výpočet parametru předurčeného zisku jsou upraveny pro výpočet parametru předurčeného zisku pouze pro frekvenční složky ležící v pásmu propustnosti filtru.

- 5 7. Detektor hlasu podle nároku 6, **v y z n a č u j í c í s e t í m**, že má dva filtry (450a, 450b) s příslušnými pásmy propustnosti a příslušné prostředky pro výpočet parametrů předurčeného zisku u příslušných pásem propustnosti a modifikační prostředky (403a, 403b) pro potlačení aktualizace v případě, že předurčený zisk přesahuje v obou pásmech propustnosti prahovou hodnotu.
- 10 8. Detektor hlasu podle nároku 6 nebo 7, **v y z n a č u j í c í s e t í m**, že zahrnuje prostředky (451a, 451b) pro vzorkování filtrovaného signálu/signálů.

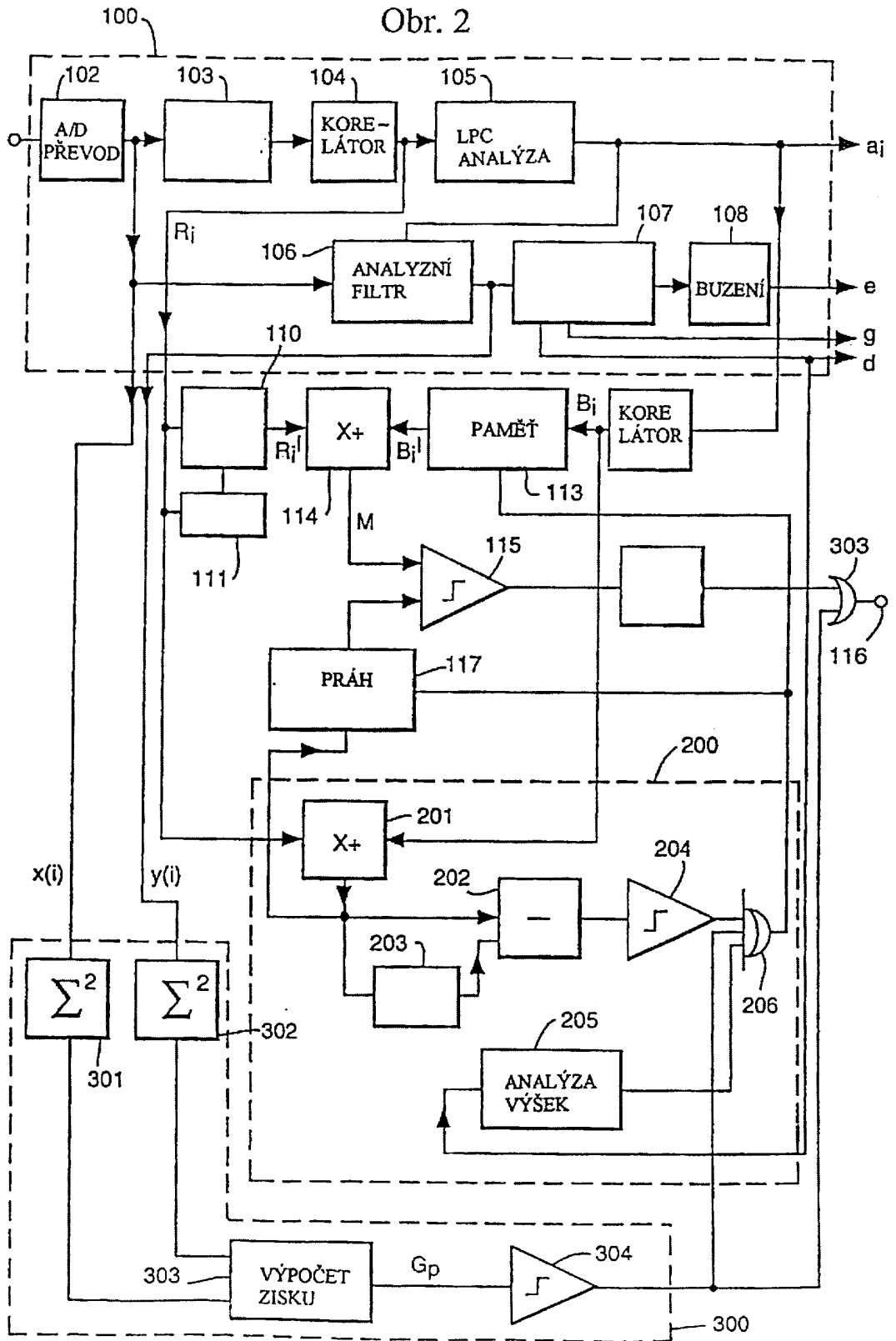
15

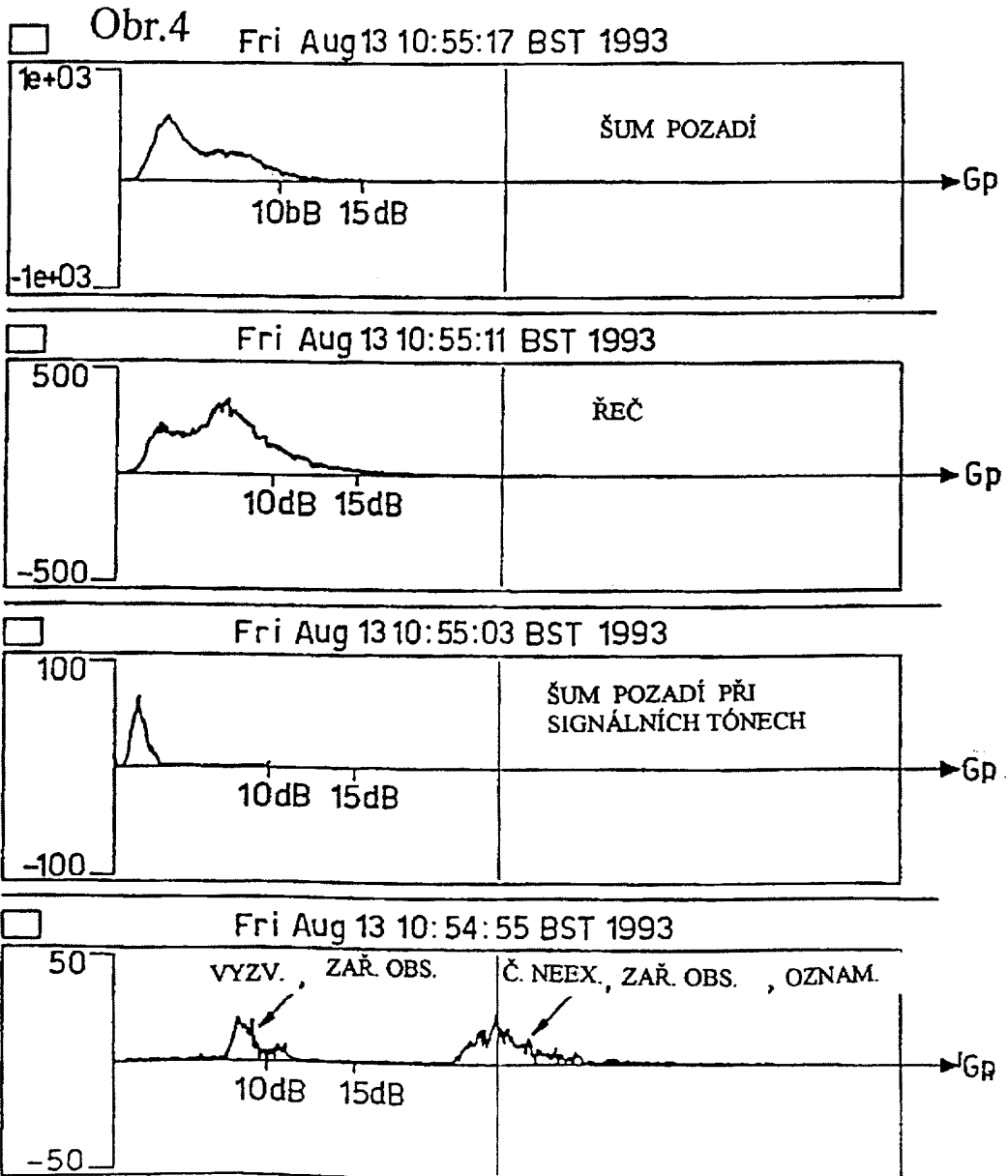
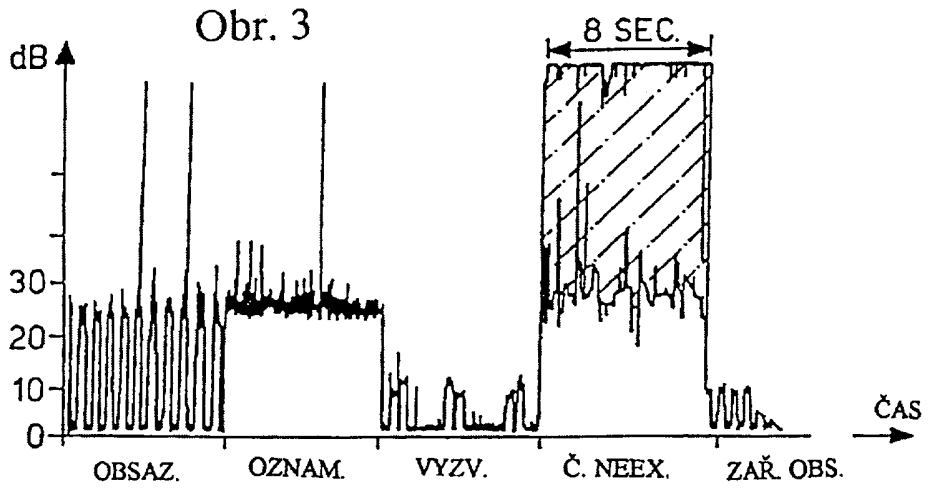
7 výkresů

Obr. 1

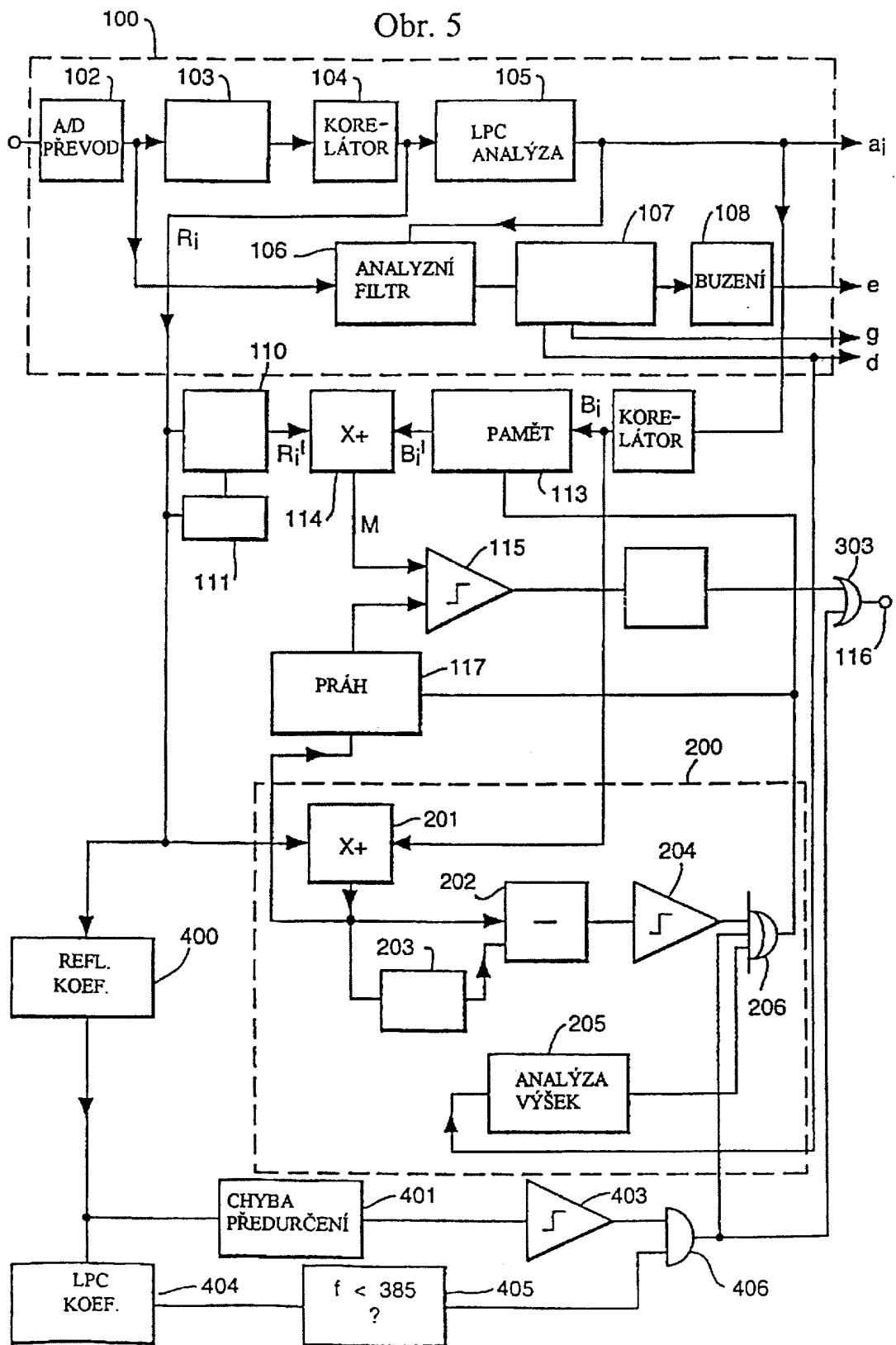


Obr. 2

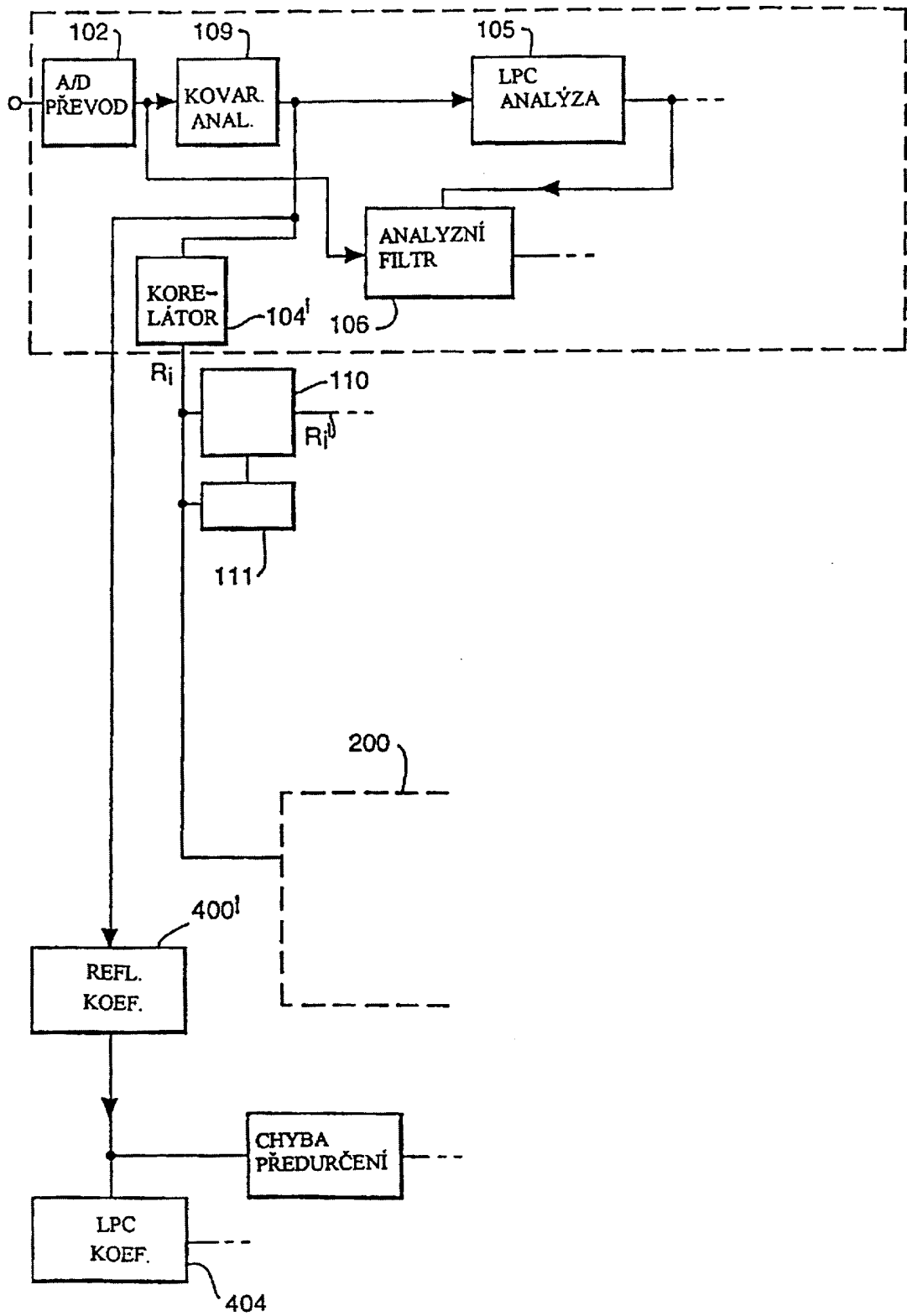




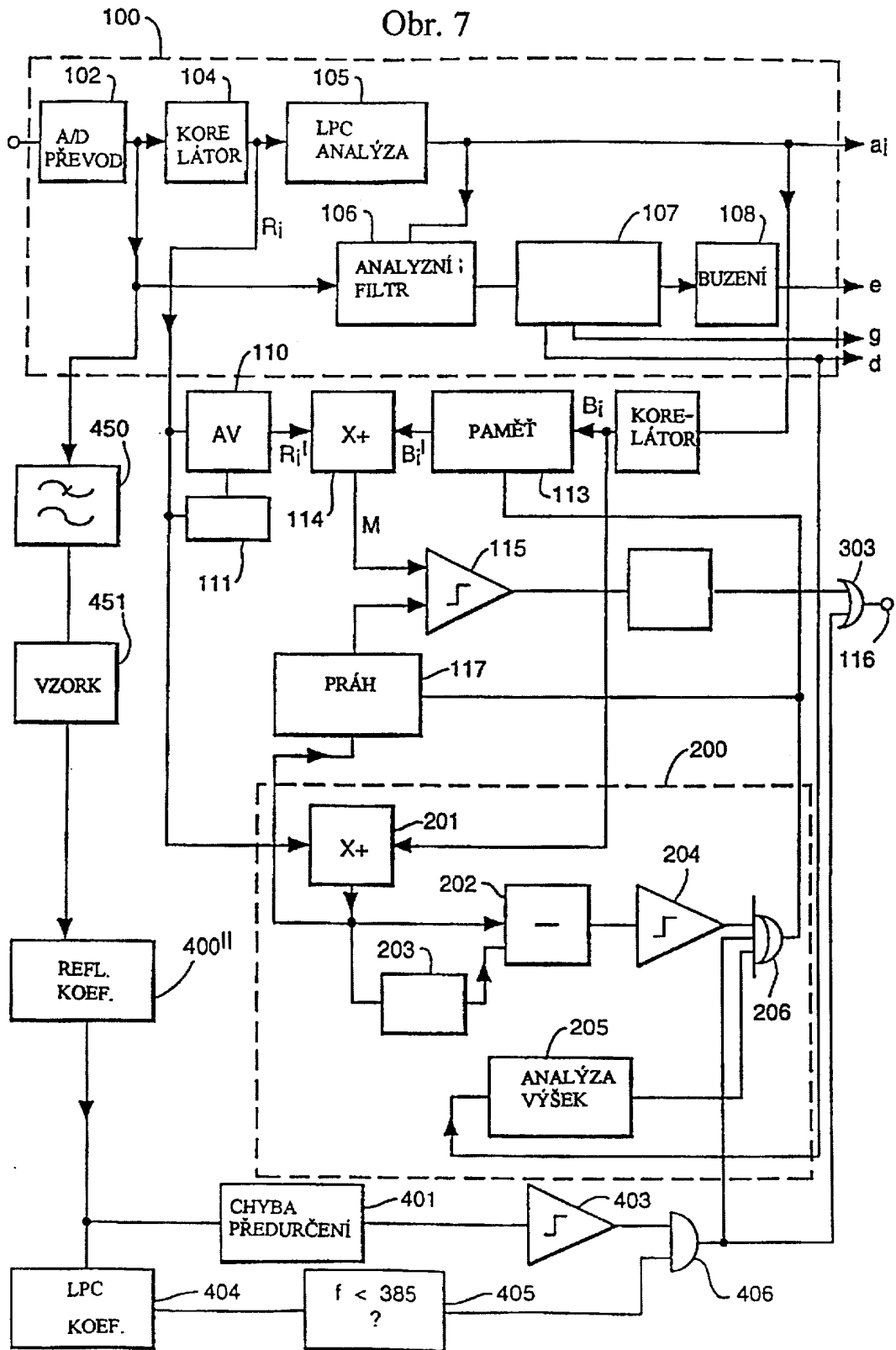
Obr. 5



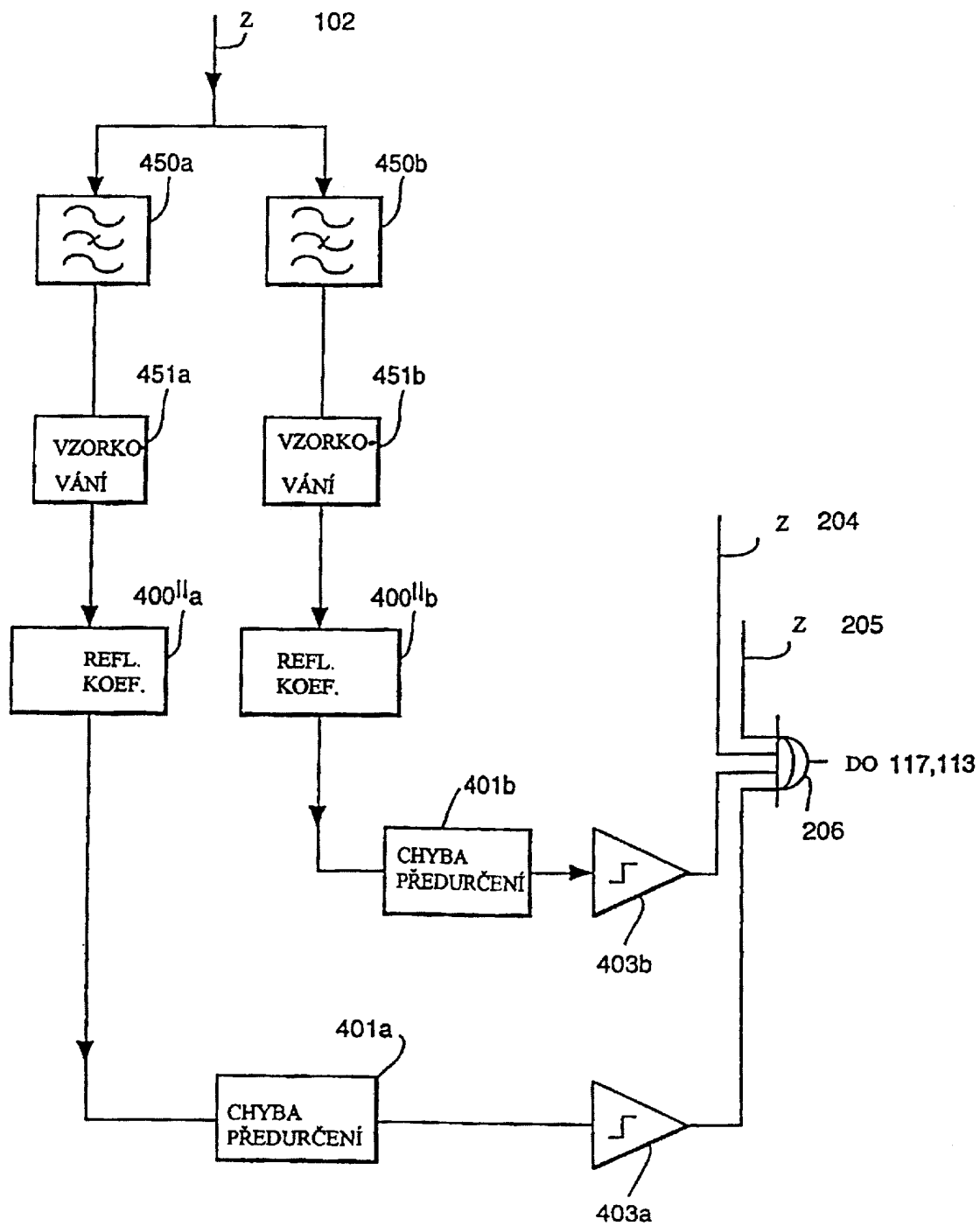
Obr. 6



Obr. 7



Obr. 8



Konec dokumentu