



ÚŘAD PRO VYNÁLEZY
A OBJEVY

POPIS VYNÁLEZU K AUTORSKÉMU OSVĚDČENÍ

(22) Přihlášeno 02 03 81
(21) (PV 1485-81)

(40) Zveřejněno 27 08 82

(45) Vydáno 15 09 85

219763

(11)

(B1)

(51) Int. Cl.³
G 10 L 1/00

(75)

Autor vynálezu

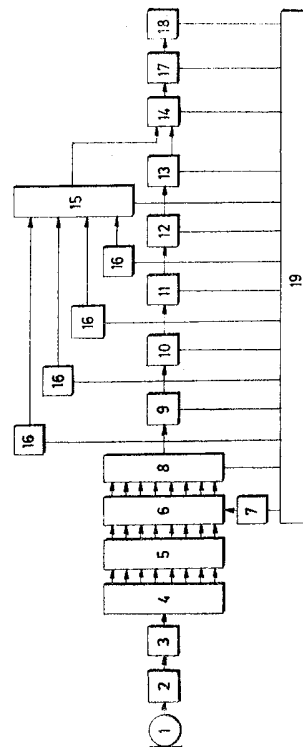
DVOŘÁK PAVEL, PRAHA, PTÁČEK MIROSLAV ing. CSc., Kladno

(54) Zařízení pro identifikaci izolovaně vyslovovaných slov s omezeným rozsahem slovníku

1

Vynález se týká zařízení pro identifikaci izolovaně vyslovovaných slov s omezeným slovníkem, které srovnává normalizovaný tvar mikrofonního signálu se stejně uspořádanými vzory, přičemž neznámému slovu přiřadí význam, odpovídající nejbližšímu vzoru slovníku. Vynález zlepšuje poznatelnost systému v hlučném prostředí, dosahuje se nezávislosti na hlasitosti mluveného slova, pokud tato překročí nadprahovou hodnotu. Ekonomická výhodnost plyne z menší postačující kapacity paměti pro uložené vzory (cena paměti tvoří podstatnou část ceny kompletního zařízení). Řešení podle vynálezu umožňuje menší spotřebu elektrické energie související s menším rozsahem potřebných paměťových obvodů.

2



Vynález se týká zařízení pro identifikaci izolovaně vyslovovaných slov s omezeným rozsahem slovníku, které srovnává normalizovaný tvar mikrofonního signálu se stejně uspořádanými vzory, přičemž neznámému slovu přiřadí význam, odpovídající nejbližšímu vzoru slovníku.

Dosud známá zařízení pro identifikaci izolovaně vyslovovaných slov, která srovnávají normalizovaný tvar mikrofonního signálu se stejně uspořádanými vzory, jsou s dobrou poznatelností (to je alespoň 95 %) použitelná pro rozsah slovníku 20 — 50 povelů (podle výběru slov a podle hlasového projevu mluvčího, na kterého nebo na které je zařízení adaptováno). Přitom používají poměrně složitá vyhodnocovací kritéria a potřebují i značnou kapacitu paměti pro uložení vzoru, nejméně 512 bitů/vzor. Některá z nich, užívající pomalejší (tj. levnější) centrální jednotky, jsou navíc pomalá. Chybovost známých systémů prudce vzrůstá při zvětšujícím se hluku okolí, který může být vyhodnocen jako mluvené slovo.

Výše uvedené nedostatky jsou zmírněny zařízením pro identifikaci izolovaně vyslovovaných slov s omezeným rozsahem slovníku podle vynálezu, jehož podstatou je takové provedení, že k mikrofonu jsou v kaskádě připojeny omezovač podprahového signálu, lineární omezovač, osmikanálový hřebenový filtr, osmikanálový usměrňovač, osmikanálový integrátor s paralelně připojeným vybíjecím obvodem spojeným s řídicí jednotkou, osmikanálový kruhový přepínač, spojený s řídicí jednotkou, logaritmický analogově-číslcový převodník, spojený s řídicí jednotkou, prahový obvod, spojený s řídicí jednotkou, osmibitový extraktor, spojený s řídicí jednotkou, šedesátibitový normalizátor, spojený s řídicí jednotkou a šedesátibitová paměť slova s řídicí jednotkou, a že na jeden vstup Hammingova komparátoru, spojeného s řídicí jednotkou je připojena šedesátibitová paměť slova a na druhý vstup Hammingova komparátoru je připojen výstup šedesátibitového kruhového přepínače, spojeného s řídicí jednotkou, na jehož vstupy jsou připojeny alespoň dvě šedesátibitové paměti vzorů, spojené s řídicí jednotkou a že na výstup Hammingova komparátoru je připojen vstup klasifikátoru, spojeného s řídicí jednotkou, na jehož výstupu je připojen výstupní obvod, spojený s řídicí jednotkou. Omezovač podprahového signálu a lineární omezovač může tvořit jediný obvod. Prahový obvod, osmibitový extraktor, šedesátibitový normalizátor, šedesátibitová paměť slova, Hammingův komparátor, šedesátibitový kruhový přepínač, šedesátibitová paměť vzorů, klasifikátor, výstupní obvod a řídicí jednotka mohou být realizovány jediným mikroprocesorovým systémem nebo jiným počítačem.

Zařazením omezovače podprahového sig-

nálu a lineárního omezovače v kaskádě za mikrofonem se zlepši poznatelnost systému v hlučném prostředí a dosáhne se nezávislosti na hlasitosti mluveného slova, pokud tato překročí padprahovou hodnotu. Při použití osmibitového extraktoru, šedesátibitového normalizátoru, šedesátibitové paměti slova a šedesátibitových pamětí vzorů podle vynálezu se pak při rozsahu slovníku 20 — 30 povelů (podle výběru slov a podle hlasového projevu mluvčího) dosáhne 95 % poznatelnosti se zařízením ekonomičtěji výhodnějším (lacnějším). Ekonomická výhodnost plyne z menší kapacity paměti pro uložení vzorů (64 bitů na vzor místo 512 bitů a více) a ze skutečnosti, že cena paměti tvoří podstatnou část ceny kompletního zařízení. Řešení podle vynálezu umožňuje menší spotřebu elektrické energie, související s menším rozsahem potřebných paměťových obvodů. Energetická úspora zvláště vynikne při aplikaci mikroprocesoru.

Na připojeném obr. 1 je znázorněn příklad zapojení zařízení pro identifikaci izolovaně mluvených slov s omezeným rozsahem slovníku. Za mikrofonem **1** jsou kaskádě zapojeny omezovač podprahového signálu **2**, lineární omezovač **3**, osmikanálový hřebenový filtr **4**, osmikanálový usměrňovač **5**, osmikanálový integrátor **6** s paralelně připojeným vybíjecím obvodem **7**, osmikanálový kruhový přepínač **8**, logaritmický analogově-číslcový převodník **9**, prahový obvod **10**, osmibitový extraktor **11**, šedesátibitový normalizátor **12** a šedesátibitová paměť slova **13**. Na jeden vstup Hammingova komparátoru **14** je připojena šedesátibitová paměť slova **13** a na druhý vstup Hammingova komparátoru **14** je připojen výstup šedesátibitového kruhového přepínače **15**, na jehož vstupy jsou připojeny alespoň dvě, na obr. 1 čtyři, šedesátibitové paměti vzorů **16**. Na výstup Hammingova komparátoru **14** je připojen vstup klasifikátoru **17** a na jeho výstup je připojen výstupní obvod **18**. Řídicí jednotka **19** je spojena s vybíjecím obvodem **7**, s osmikanálovým kruhovým přepínačem **8**, s logaritmickým analogově-číslcovým převodníkem **9**, s prahovým obvodem **10**, s osmibitovým extraktorem **11**, s šedesátibitovým normalizátorem **12**, s šedesátibitovou pamětí slova **13**, s Hammingovým komparátorem **14**, s šedesátibitovým kruhovým přepínačem **15**, s šedesátibitovými pamětmi slova **16**, s klasifikátorem **17** a s výstupním obvodem **18**.

Funkce jednotlivých bloků znázorněných na obr. 1 je dále podrobněji vysvětlena. Mikrofon **1** snímá akustický signál mluveného slova a převádí ho na signál elektrický. Omezovač podprahového signálu **2** pracuje jako expander dynamicky s nastavitelným prahem (tvarem expanzní charakteristiky); omezuje zesílení podprahových signálů. Práh

se nastaví v závislosti na hluku prostředí, z něhož je mluvené slovo snímáno tak, aby byl rušivý signál potlačen. Lineární omezovač **3** pracuje jako lineární zesilovač až do maximálního přípustného signálu, odpovídajícího rozsahu logaritmického analogově-číslicového převodníku **9**. Po překročení tohoto signálu se zesílení lineárního omezovače **3** progresivně zmenšuje. Osmikanálový hřebenový filtr **4** je tvořen osmi pásmovými propustmi, jejichž propouštěná pásma na sebe navazují a zahrnují pásmo kmitočtů, odpovídající pásmu použitého mikrofonu **1**. Osmikanálový usměrňovač **5** může usměrňovat jednu nebo obě půlvlny vstupních signálů. Osmikanálový integrátor **6** má nabíjecí časovou konstantu 3 — 15 ms, optimálně 5 ms a vybíjecí časovou konstantu 20 — 150 ms, optimálně 50 ms. Vybíjecí obvod **7** vybíjí integrační prvky osmikanálového integrátoru **6** s časovou konstantou 0 — 2 ms (optimálně 0,5 ms). Osmikanálový kruhový přepínač **8** přepíná signály osmi vstupů na jediný výstup s kruhovou frekvencí 30 — 80 Hz (optimálně 50 Hz), přičemž je všech osmi vstupů přepnuto včetně následujícího vybití integračních prvků v době menší, než 1/10 periody kruhové frekvence; ve zbylé části periody je signál zpracován osmibitovým extraktorem **11**. Logaritmický analogově-číslicový převodník **9** převádí vstupní analogový signál o dynamickém rozpětí 20 — 80 dB, optimálně 64 dB, na 2 — 16 bitový, optimálně 4 bitový číslicový údaj. Prahový obvod **10** nahradí podprahové číslicové údaje pulsy. Při aplikaci 4 bitového logaritmického analogově-číslicového převodníku **9** je vhodné nastavení prahu 2 — 6, optimálně 4 v závislosti na špičkových úrovních rušivého hluku prostředí, z něhož je mluvené slovo snímáno. Osmibitový extraktor **11** nalezne ve sledu osmi čísel, odpovídajících úrovním výstupních signálů osmikanálového hřebenového filtru **4** z jedné periody osmikanálového kruhového přepínače **8** čísla, která jsou ve vztahu k ostatním relativními spektrálními maximy a nahradí v nich původní hodnotu číslem **1**. Ostatní čísla osmice osmibitový extraktor **11** vynuluje, stejně tak, jako čísla tvořící kontury mezi sousedícími plochými relativními maximy, zabírajícími více frekvenčních pásem. Takto vzniklou osmicí jednobitových čísel osmibitový extraktor **11** nakonec shrne do jediného osmibitového čísla. Šedesátičtyřbitový normalizátor **12** počne takto vzniklá čísla, od prvního nenulového čísla, ukládat do paměti jako základní popis mluveného slova. Ukládání se ukončí po dosažení zadaného počtu čísel, optimálně 50 nebo po uložení více čísel, než odpovídá zadané minimální délce slova, optimálně 20, za níž následuje zadaný počet pulsových čísel, optimálně 16, odpovídajících nulovému signálu za slovem. Krátké poruchy, menší než zadaná délka slova, optimálně

20 a vzdálené od nenulového signálu alespoň o zadaný počet nulových čísel, optimálně 16, šedesátičtyřbitový normalizátor **12** ignoruje. Po uložení základního popisu mluveného slova do paměti šedesátičtyřbitový normalizátor **12** vymezí délku mluveného slova, rozdělí ji do osmi stejně dlouhých časových úseků a jednotlivé, sobě si odpovídající bity všech čísel základního popisu mluveného slova v každém úseku zprůměruje. Vzniklé šedesátičtyřbitové číslo uloží šedesátičtyřbitový normalizátor **12** do šedesátičtyřbitové paměti slova **13**. Šedesátičtyřbitový kruhový přepínač **15** pak přepíná šedesátičtyřbitové paměti vzorů **16**, pro jedno slovo může být libovolný počet vzorů, na druhý vstup Hammingova komparátoru **14**, jehož první vstup je připojen na šedesátičtyřbitovou paměť slova **13**, Hammingův komparátor **14** určí Hammingovu vzdálenost mezi dvěma šedesátičtyřbitovými čísly na svých vstupech. Vypočtená vzdálenost spolu s údajem o čísle vzoru jde z výstupu Hammingova komparátoru **14** na vstup klasifikátoru **17**. Zde se určí číslo nejbližšího vzoru nebo jemu odpovídajícího slova a vyšle se z výstupu klasifikátoru **17** na vstup výstupního obvodu **18**, který číslo identifikovaného slova předá do dalšího spolupracujícího nebo řízeného zařízení, které již není předmětem vynálezu. Ze spolupracujícího zařízení lze přijímat i příkaz pro řídicí jednotku **19**, která činnost zařízení pro identifikaci izolovaných slov s omezeným rozsahem slovníku koordinuje, k sejmutí a k identifikaci mluveného slova.

Zařízení pro identifikaci izolovaně vyslovovaných slov s omezeným rozsahem slovníku podle obr. 1 lze programovou úpravou realizovat jako adaptibilní, a to na nový slovník nebo na nového mluvčího nebo skupinu mluvčích. Při adaptaci se obsahy šedesátičtyřbitové paměti slova **13**, odpovídající jednotlivým slovům nebo jejich odpovídajícím realizacím, přepisují prostřednictvím řídicí jednotky **19** do šedesátičtyřbitových pamětí vzorů **16**, a to s údaji o číslech slov, jímž jednotlivé šedesátičtyřbitové paměti vzorů **16** odpovídají. Zařízení pro identifikaci izolovaně vyslovovaných slov s omezeným rozsahem lze realizovat i tak, že omezovač podprahového signálu **2** a lineární omezovač **3** tvoří jediný celek řízený zesilovačem, nebo i tak, že prahový obvod **10**, osmibitový extraktor **11**, šedesátičtyřbitový normalizátor **12**, šedesátičtyřbitová paměť slova **13**, Hammingův komparátor **14**, šedesátičtyřbitový kruhový přepínač **15**, šedesátičtyřbitové paměti vzorů **16**, klasifikátor **17** a řídicí jednotka **19** jsou realizovány buď všechny, nebo některé z nich, jediným mikroprocesorovým systémem nebo jiným počítačem. Funkci klasifikátoru **17** lze doplnit ještě odmítnutím identifikace v případě, kdy Hammingova vzdálenost mezi obsahem še-

desátičtyřbitové paměti slova 13 a jemu nejbližšího obsahu šedesátičtyřbitové paměti

vzoru 16 překročí zadanou hodnotu optimálně 21.

PŘEDMĚT VYNÁLEZU

1. Zařízení pro identifikaci izolovaně vyslovovaných slov s omezeným rozsahem slovníku, vyznačené tím, že k mikrofonu (1) jsou v kaskádě připojeny omezovač podprahového signálu (2), lineární omezovač (3), osmikanálový hřebenový filtr (4), osmikanálový usměrňovač (5), osmikanálový integrátor (6) s paralelně připojeným vybíjecím obvodem (7) spojeným s řídicí jednotkou (19), dále osmikanálový kruhový přepínač (8) spojený s řídicí jednotkou (19), logaritmický analogově číslicový převodník (9) spojený s řídicí jednotkou (19), prahový obvod (10) spojený s řídicí jednotkou (19), osmibitový extraktor (11) spojený s řídicí jednotkou (19), šedesátičtyřbitový normalizátor (12) spojený s řídicí jednotkou (19) a šedesátičtyřbitová paměť slova (13) spojená s řídicí jednotkou (19), přičemž na je-

den vstup Hammingova komparátoru (14) spojeného s řídicí jednotkou (19) je připojena šedesátičtyřbitová paměť slova (13) a na druhý vstup Hammingova komparátoru (14) je připojen výstup šedesátičtyřbitového kruhového přepínače (15) spojeného s řídicí jednotkou (19), na jehož vstupy jsou připojeny alespoň dvě šedesátičtyřbitové paměti vzoru (16) spojené s řídicí jednotkou (19), přičemž na výstup Hammingova komparátoru (14) je připojen vstup klasifikátoru (17) spojeného s řídicí jednotkou (19), na jehož výstup je připojen výstupní obvod (18) spojený s řídicí jednotkou (19).

2. Zařízení pro identifikaci izolovaně vyslovovaných slov s omezeným rozsahem slovníku podle bodu 1, vyznačené tím, že omezovač podprahového signálu (2) a lineární omezovač (3) tvoří jediný obvod.

1 list výkresů

