

(19) REPUBLIKA SRBIJA

(12) Patentni spis

(11) 49857B



(51) Int. Cl.⁷

H 03 F 3/00

ZAVOD ZA
INTELEKTUALNU SVOJINU
BEOGRAD

(21) Broj prijave: **P- 2006/0611**

(22) Datum podnošenja prijave: **03.11.2006.**

(43) Datum objavljivanja prijave: **04.06.2007.**

(45) Datum objavljivanja patenta: **07.08.2008.**

(30) Međunarodno pravo prvenstva:

(61) Dopunski patent uz osnovni
patent broj:

(62) Izdvojen patent iz prvobitne
prijave broj:

(73) Nosilac patenta:

**MicronasNIT,
Fruškogorska 11a, 21000 Novi Sad, RS**

(72) Pronalazači:

Šarić dr Z.;
Jovičić dr S.;
Kovačević dr V.;
Teslić dr N.;
Pap dr I.

(74) Zastupnik:

(54) Naziv: **POSTUPAK I SISTEM ZA AUTOMATSKU
REGULACIJU POJAČANJA (AGC) NA OSNOVU
OČITAVANJA MIKROFONSKOG NIZA**

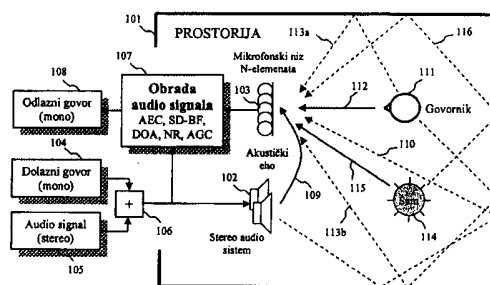
(51) Int. Cl.⁷

H 03 F 3/00

(57) Apstrakt:

Pronalazak se odnosi na postupak i sistem za automatsku kontrolu pojačanja (AGC) govornog signala snimljenog pomoću mikrofonskog niza u specifičnim "handsfree" telekonferencijskim uslovima komuniciranja i u akustičkom ambijentu sa velikim nivoom raznih smetnji kao što su: akustički eho, ambijentalni šum, reverberacija i drugi govornici, kao i različitog položaja aktuelnog govornika u prostoru u odnosu na mikrofonski niz. Postupak za AGC se sastoji od: estimacije trajektorije vršne snage govornog signala, estimacije snaga akustičkih smetnji i njihovog participiranja u određivanju koeficijenta pojačanja, određivanja nagiba karakteristike kompresije, određivanja koeficijenta pojačanja za AGC i formiranja izlaznog govornog signala sa minimalnim distorzijama. Sistem sadrži jedan ulaz i jedan izlaz za aktuelni govorni signal, više ulaza na koje se dovode signali za estimaciju akustičkih smetnji, i tri ulaza na kojima se postavljaju tri konstante koje definišu polje aktivnog

funkcionisanja AGC: nominalna snaga izlaznog signala, dinamika promene pojačanja AGC sistema i maksimalni nagib karakteristike kompresije.



Slika 1.

RS 49857 B

OBLAST TEHNIKE NA KOJU SE PRONALAZAK ODNOSI

Pronalazak pripada oblasti obrade akustičkog signala, ili konkretnije, metodama automatske kontrole pojačanja (AGC) govornog signala snimljenog pomoću mikrofonskog niza u specifičnim „hands-free“ telekonferencijskim uslovima komuniciranja.

TEHNIČKI PROBLEM

Slobodni, „hands-free“ (engl.), komunikacioni sistemi za prenos govornog signala u punom dupleksu koriste se u mnogim aplikacijama kao što su: video-telefonski sistemi, telekonferencijski sistemi, spikerfoni u prostoriji ili kolima, komunikacija čovek-računar putem glasa, itd. „Hands-free“ govorna komunikacija podrazumeva da se govornik nalazi u akustičkom ambijentu na određenoj distanci od interfejsnih elemenata sistema - mikrofona i zvučnika. Ovakvi uslovi odvijanja govorne komunikacije generišu više tehničkih problema koje je potrebno rešiti da bi se održao kvalitet komunikacije na prihvatljivom nivou. Jedan od problema je i održavanje nivoa predajnog govornog signala na približno konstantnom nivou nezavisno od tipa govornika ili njegove udaljenosti od mikrofona.

Naime, tri činjenice su ključne u prethodno naznačenom problemu. Prvo, intenzitet glasa različitih govornika zavisi od njihovog fonetskog kapaciteta. To je prirodna individualna karakteristika svakog čoveka, ali može biti i stečena karakteristika u zavisnosti od socijalnih ili habitualnih faktora. Na primer, govornik iz ruralnih krajeva najčešće ima povišen intenzitet govora u odnosu na govornika iz gradske sredine. Drugo, kod upotrebe „hands-free“ komunikacionih sistema govornik može biti na različitim rastojanjima od mikrofona. Poznato je da intenzitet zvuka opada sa kvadratom rastojanja što u prethodno pomenutom slučaju unosi znatne varijacije u veličini govornog signala na izlazu mikrofona. I treće, posebno kod telekonferencijskih

aplikacija, govornik je često u pokretu i menja lokaciju u prostoru u odnosu na položaj mikrofona, što prouzrokuje variranje intenziteta govornog signala na izlazu mikrofona.

Pored navedenih tehničkih problema pojavljuje se i problem variranja intenziteta glasa u slučajevima kada su u prostoriji prisutni ambijentalna buka, reverberacija ili drugi govornici, što je tipičan slučaj kod „cocktail-party” efekta, zbog čega aktuelni govornik podiže nivo svog glasa (tzv. Lombardov efekat). Svi navedeni tehnički problemi rešavaju se tehničkim rešenjem pod poznatim nazivom „automatska kontrola pojačanja” - AGC (*Automatic Gain Control*, engl.). AGC kao tehničko rešenje podešava pojačanje aktuelnog govornog signala održavajući približno konstantnim njegov nivo na svom izlazu. Kada je govorni signal u pitanju ova kontrola se obavlja u oba domena: amplitudskom i vremenskom.

Dodatna dva problema AGC kao tehničko rešenje mora da reši u slučaju telekonferencijskih primena, a to su: a) diskriminacija govorni signal – ambijentalne smetnje, gde treba obuhvatiti signale reverberacije i kompetitivne govornike, i b) brzina reakcije AGC rešenja na pojavu ili nestanak govornog signala u analiziranom signalu, što zahteva dobru usklađenost rada AGC rešenja sa prirodom govornog signala.

Rešenje svih navedenih problema treba da ima za cilj obezbeđivanje maksimalne razumljivosti, prirodnosti i prijatnosti slušanja govornog signala u kompleksnim telekonferencijskim uslovima primene AGC rešenja.

STANJE TEHNIKE

Postoje dve generalne tehnike u realizaciji AGC problema. U prvoj se koristi ulazni signal za generisanje koeficijenta pojačanja koji se množi sa ulaznim signalom i tako produkuje izlazni signal približno konstantne snage. Ova tehnika se naziva “feed forward” tehnika. Druga tehnika je tzv. “feedback” tehnika u kojoj se za generisanje koeficijenta pojačanja koristi izlazni signal, koji se poredi sa referentnim signalom (koji definiše željeni nivo izlaznog signala) a signal greške definiše koeficijent pojačanja.

Konvencionalni AGC sistemi, zasnovani na bilo kojoj od ove dve tehnike, rešavaju pitanje automatske kontrole nivoa govornog signala na bazi kompresije amplitude govornog signala koristeći nelinearnu estimaciju snage govornog signala i elemente detektora aktivnosti govora (VAD – *voice activity detector*) ili tzv. *squelch* sistem. Na primer: U.S. patent 4,947,133, prijavljen 19. januara 1988, sa naslovom “Method and apparatus for automatic signal level adjustment”, koristi “feed forward”

tehniku u adaptivnoj kompresiji govornog signala, zatim U.S. patent 5,854,845, prijavljen 21. februara 1996, sa naslovom "Method and circuit for voice automatic gain control", daje metod za automatsku kontrolu pojačanja signala uz minimiziranje distorzija na granicama govor-pauza, kao i U.S. patent 6,959,082, prijavljen 10. jula 2001, sa naslovom "Method and system for automatic gain control with adaptive table lookup", koji koristi obe tehnike "feed forward" i "feedback" u automatskoj kontroli pojačanja pri čemu se pojačanje tabelarno određuje na bazi estimirane snage ulaznog signala.

Nešto drugačiji zahtevi u realizaciji AGC nastaju u specifičnim „hands-free“ telekonferencijskim uslovima komuniciranja u punom dupleksu. Kvalitetno snimanje govora u uslovima prisustva akustičkih smetnji i reverberacije prostorije predstavlja složen problem. U uslovima kada se spektri korisnog govornog signala preklapaju sa spektrima prisutnih smetnji, jednokanalnim postupcima obrade signala nije moguće ostvariti značajnije poboljšanje kvaliteta govornog signala već se koriste višemikrofonski postupci (mikrofonski nizovi). Prednost mikrofonskih nizova u odnosu na jednokanalne postupke obrade je njihova sposobnost da prilagode svoju prostornu karakteristiku prijema (karakteristiku usmerenosti) trenutnom prostornom rasporedu odabranog govornika i smetnji. Pri tome ostvaruju maksimalno potiskivanje prisutnih smetnji uz istovremeno isticanje odabranog govornika. Osnovni problemi koji se u primeni mikrofonskih nizova sreću su sledeći (M.S. Brandstein, D.B. Ward (Eds.), *Microphone Arrays: Signal Processing Techniques and Applications*, Springer, Berlin 2001; Y. Huang, J. Benesty, *Audio signal processing for next generation multimedia communication systems*, Kluwer Academic Publishers Publ., 2004.): nepoznavanje tačne lokacije odabranog govornika, nepoznavanje broja i prostornog rasporeda prisutnih smetnji, višestruke refleksije korisnog izvora i smetnji o zidove prostorije i nestacionarnost izvora akustičkih smetnji i odabranog govornika.

Kada se mikrofonski niz upotrebi u telekonferencijskim sistemima koji funkcionišu u punom dupleksu, onda se broj problema uvećava. Najveći problem je pojava akustičkog eha, zatim dvostruka govorna aktivnost (aktivnost u oba smera komuniciranja), kao i moguća pojava nestabilnosti sistema, tzv. mikrofonija. Funkcionisanje AGC je znatno komplikovanije u ovakvim uslovima (K. Kobayashi, Y. Haneda, K. Furuya, A. Kataoka, "A hands-free unit with adaptive microphone array for directional AGC", *2005 International Workshop on Acoustic Echo and Noise Control*, Eindhoven, The Netherlands, September 12 - 15, 2005).

Integralno rešenje AGC, izloženo u ovom patentu, objedinjuje pozitivne osobine nekih konvencionalnih rešenja sa novim rešenjima u multimikrofonskom okruženju i zahtevima koje nameće „hands-free“ telekonferencijska aplikacija u punom dupleksu obezbeđujući kvalitetnu govornu komunikaciju.

IZLAGANJE SUŠTINE PRONALASKA

Predmet ovog pronalaska je metod i sistem za automatsku kontrolu pojačanja (AGC) govornog signala snimljenog pomoću mikrofonskog niza u složenom akustičkom ambijentu i u specifičnim „hands-free“ telekonferencijskim uslovima komuniciranja a u cilju obezbeđenja kvaliteta i razumljivosti govornog signala.

Sušтина pronalaska jeste u specifičnoj obradi govornog signala koji se snima u akustičkom ambijentu prostorije u kojoj se nalazi sistem i govornik. Za snimanje govornika u prostoriji, koji se nalazi na određenom rastojanju (do nekoliko metara), koristi se mikrofonski niz od N mikrofona. Mikrofonski niz snima sve signale u prostoriji: koristan signal kao direktan talas koji stiže od govornika do mikrofona i signale smetnji koji mogu biti raznovrsni. Kao signali smetnje pojavljuju se: akustički eho kao direktan zvučni talas iz zvučnika preko kojih se emituje glas sagovornika sa udaljenog kraja komunikacionog kanala, direktni talasi od jednog ili više izvora šumova ili izvora drugih smetnji koji se mogu naći u prostoriji i svi reflektovani talasi (eho prostorije) koji potiču od svih izvora zvukova, uključujući i govornika, a koji nastaju usled reverberacije prostorije. Treba naglasiti da izvori zvukova u prostoriji mogu biti stacionarni ili nestacionarni, što je najčešći slučaj, kako po svojim karakteristikama tako i po lokaciji u prostoriji (pokretni izvori zvukova). Suština pronalaska jeste da se očuva kvalitet i razumljivost govornog signala aktuelnog govornika uprkos velikim ambijentalnim smetnjama, različitim mogućim prostornim položajima govornika u odnosu na mikrofonski niz i, konsekventno, velikim varijacijama u intenzitetu govornog signala.

Specifičan aspekt pronalaska se nalazi u kompletnoj obradi signala u frekvencijskom domenu, koji omogućava određene prednosti u pogledu brzine obrade i broja računskih operacija, što je veoma važno za implementaciju u realnom vremenu.

Sledeću specifičnost pronalaska čini nelinearna karakteristika amplitudske kompresije. Nagib ove karakteristike je određen na kompleksan način preko trajektorije vršne snage govornog signala. Analiza obuhvata trend i konveksnost trajektorije, i

relativno poređenje ovih parametara sa trenutnom vršnom snagom govornog signala. Dakle, u potpunosti su uzete u obzir karakteristike govornog signala, te je karakteristika kompresije optimizirana za ovu vrstu signala.

Određivanje koeficijenta pojačanja je naredna specifičnost pronalaska. On se određuje na bazi nagiba karakteristike kompresije kao i na bazi veličine vršne snage ulaznog govornog signala i srednjih snaga smetnji koje se pojavljuju u sistemu slobodne govorne komunikacije na bazi mikrofonskog niza.

Sledeću specifičnost pronalaska čini postupak adaptivnog određivanja koeficijenta pojačanja na bazi detektovanih informacija o pauzama u govornom signalu, prisutnosti rezidualnog eha u sistemu i prisutnosti konkurentnog govornika ili akustičke smetnje. Ove informacije omogućavaju pravilno funkcionisanje AGC kod različitog sadržaja ulaznog signala i obezbeđuju poboljšanje kvaliteta izlaznog signala.

Posebnu specifičnost pronalaska čini set parametara koji omogućava optimalan izbor uslova rada AGC u skladu sa zahtevima aplikacije. Ovi parametri definišu nominalni nivo izlaznog signala, maksimalan nagib karakteristike kompresije i maksimalno pojačanje AGC. Na taj način je postignuta fleksibilnost rešenja.

Inventivnost u ovom pronalasku se nalazi u poboljšanju svake od navedenih specifičnosti a posebno u činjenici da su na posredan način u algoritam estimacije koeficijenta pojačanja unete informacije o ambijentalnim smetnjama kao i informacija o aktivnosti aktuelnog govora.

Ovi i drugi aspekti, specifičnosti i benefiti ovog pronalaska biće očigledniji nakon uvida u detaljan opis pronalaska, patentne zahteve i pripadajuće crteže.

KRATAK OPIS SLIKA I NACRTA

Slika 1 – prikazuje ambijentalne uslove primene sistema za slobodnu video-telefonsku komunikaciju pomoću mikrofonskog niza.

Slika 2 – prikazuje blok dijagram sistema za obradu audio signala u okviru sistema za slobodnu video-telefonsku komunikaciju i lociranje modula za AGC; osnovni blokovi ovog sistema su: blok za potiskivanje eha (AEC), blok za formiranje karakteristike usmerenosti mikrofonskog niza (SD-BF), blok za lociranje govornika u prostoru (DOA), blok za potiskivanje šuma (NR) i blok za automatsku kontrolu pojačanja (AGC).

Slika 3 - prikazuje postupak formiranja prenosne karakteristike AGC.

Slika 4 – prikazuje blok dijagram osnovne strukture sistema automatske kontrole pojačanja (AGC).

Slika 5 – prikazuje tok operacija u izračunavanju faktora pojačanja A_{agc} .

Slika 6 – prikazuje interfejsne tačke i signale AGC sistema.

DETALJAN OPIS PRONALASKA

Ovaj pronalazak opisuje sistem i postupak automatske kontrole pojačanja (AGC) govornog signala u sistemu slobodne govorne komunikacije pomoću mikrofonskog niza. Da bi se razumela potreba i značaj AGC u ovakvim sistemima na slici 1 prikazani su ambijentalni uslovi odvijanja slobodne govorne komunikacije, dok je na slici 2 prikazan blok dijagram sistema za obradu audio signala u okviru kojeg je lociran modul za AGC govornog signala.

Slika 1 šematski prikazuje ambijentalne uslove primene sistema za slobodnu video-telefonsku komunikaciju pomoću mikrofonskog niza. U prostoriji 101 nalaze se sistem za slobodnu video-telefonsku komunikaciju, govornik 111 i izvor šuma 114, što je uobičajeno za svaki akustički ambijent. Preko zvučnika 102 stereo audio sistema govornik 111 sluša dolazni govorni signal 104 sagovornika sa udaljenog kraja najčešće kao mono signal; moguće je da se preko stereo audio sistema emituje i drugi audio signal. Zvuk u ambijentu prostorije 101 snima mikrofonski niz 103 sastavljen od N mikrofona. Nakon kompleksne obrade mikrofonskih signala u bloku 107 govorni signal govornika 111 se preko bloka 108 prenosi ka udaljenom sagovorniku kao mono signal.

Ambijentalni uslovi odvijanja govorne komunikacije u prostoriji 101 su veoma kompleksni. Kod slobodne video-telefonske komunikacije u prostoriji 101 postoji minimum tri izvora zvuka: stereo zvučnici 102 koji emituju govor udaljenog sagovornika i drugi audio signal, govornik 111 i bar jedan izvor šuma 114. U prostoriji može biti i više izvora šumova: šum računara, šum klima sistema, buka sa ulice koja prodire u prostoriju kroz prozore, buka iz susednih prostorija, vibracije zgrade, ili drugi govornik, više govornika, izvor muzike, itd. Dakle, pojavljuje se veoma složena akustička slika u prostoriji. Mikrofonski niz 103 snima, kao senzorski sistem, sve zvuke u prostoriji, snima direktne zvučne talase od svakog izvora ali i sve refleksije od zidova prostorije i drugih predmeta koji se nalaze u njoj. Tako na primer, od zvučnika 102 do

mikrofonskog niza 103 stiže direktan talas 109 i mnogi reflektovani talasi od kojih je samo jedan 110 prikazan na slici 1; od govornika 111 stiže direktan talas 112 i pored ostalih i dva reflektovana talasa 113a i 113b, od izvora šuma 114 stiže direktan talas 115 i pored ostalih i reflektovani talas 116.

Od svih zvukova koje mikrofonski niz snima jedino je direktan talas 112 od govornika 111 koristan signal, svi ostali su smetnje. Akustički eho 109 koji dolazi iz zvučnika 102, obično predstavlja smetnju najvećeg intenziteta. Sve ostale refleksije zbirno čine reverberaciju prostorije. Zadatak bloka za obradu audio signala 107 jeste da potisne signal akustičkog eha, da selektuje koristan signal 112 od svih ostalih smetnji, da potisne signale reverberacije i da potisne direktne signale izvora smetnji, kojih može da bude i više od jednog izvora. Poseban zadatak bloka 107 jeste adaptivno praćenje nestacionarnosti akustičke scene u prostoriji bilo da se govornik pokreće, ili da se od razgovora do razgovora nalazi na različitim pozicijama u prostoriji, ili da se izvori šumova pokreću, da su nestacionarni ili da menjaju svoje karakteristike. Sa aspekta udaljenog sagovornika odlazni govorni signal 108 treba da je stabilan i nezavisan od navedenih akustičkih varijabilnosti, što treba da obezbedi kvalitetnu i prijatnu govornu komunikaciju.

Na slici 2 prikazan je blok dijagram sistema za obradu audio signala u okviru sistema za slobodnu video-telefonsku komunikaciju, koji je na slici 1 označen sa 107. Mikrofonski signali 103, od M1 do MN, kao i signali stereo zvučnika 102, Zv-L i Zv-D, se preko ulaznog interfejsa 201 uvode u blok 202, N-kanalni akustički potiskivač eha – AEC (*Acoustic Echo Cancelling*), u kome se vrši potiskivanje akustičkog eha 109, slika 1, u snimanom govornom signalu 112, slika 1. U ulaznom interfejsu 201 vrši se konverzija signala u frekvencijski domen pomoću diskretne Fourierove transformacije (DFT) i sva dalja obrada signala se vrši u frekvencijskom domenu, na nivou bloka ulaznih odmeraka. Na izlazu bloka 202 dobijaju se signali iz mikrofona sa potisnutim akustičkim ehom, S_{AECI} do S_{AECN} . Ovi signali se uvode u blok 203, superdirektivni usmerivač – SD-BF (*Superdirective Beamformer*), koji oblikuje karakteristiku usmerenosti mikrofonskog niza 103, i u blok 204, azimut DOA (*Direction Of Arrival*), koji određuje lokaciju aktuelnog govornika u horizontalnoj ravni i daje podatak o uglu azimuta θ_a . Ovaj podatak se koristi za usmeravanje karakteristike mikrofonskog niza ka aktuelnom govorniku.

Signal na izlazu bloka 203 sadrži koristan govorni signal i signal smetnji koji se sastoji od rezidualnog signala nakon potiskivanja akustičkog eha, potisnut šum ambijenta i potisnute signale reverberacije. Ovaj signal ulazi u blok 205, potiskivač šuma - NR (*Noise Reduction*), gde se vrši dodatno potiskivanje signala smetnji. Proces potiskivanja je adaptivan obzirom na nestacionarnost signala smetnji.

Finalni blok obrade signala u sistemu za slobodnu govornu komunikaciju jeste blok 206, automatska kontrola pojačanja - AGC, koji jeste predmet ovog pronalaska. U ovom bloku koristi se više informacija iz celokupnog sistema koje su važne za definisanje mogućih uslova u kojima se govorni signal može naći i gde je potrebno na odgovarajući način izvršiti njegovu amplitudsku korekciju. Na taj način se može obezbediti približno isti nivo predajnog govornog signala nezavisno od udaljenosti aktuelni govornik od mikrofonskog niza i obezbediti njegov bolji kvalitet na udaljenom kraju komunikacionog kanala. Preko izlaznog interfejsa 207 estimirani govorni signal na bližem kraju δ se prenosi kroz komunikacioni kanal ka udaljenom sagovorniku.

Važno je zapaziti da se funkcionisanje AGC u kompleksnim ambijentalnim uslovima primene mikrofonskog niza, u višekanalnim sistemima i u slobodnoj govornoj komunikaciji, veoma razlikuje od konvencionalne primene AGC u jednokanalnim sistemima i primenama.

Na slici 3 prikazan je proces formiranja prenosne karakteristike AGC sistema. Dijagrami su dati u formi $L_{out} = f(L_{in})$. U slučaju sistema bez funkcije AGC sa jediničnim pojačanjem prenosna karakteristika bi bila 301, K_1 , koja obezbeđuje samo prosleđivanje ulaza na izlaz. Osnovni zadatak AGC sistema je da govorni signal na svom izlazu održi na istom optimalnom nivou snage 302, L_{opt} , i u idealnom slučaju – kada imamo potpunu kompresiju, svaki ulazni nivo bi se dovodio na nivo L_{opt} .

Prenosna karakteristika K_2 kojom se ostvaruje potpuna kompresija dinamike ulaza nije dobra u uslovima prisustva šuma u sistemu, jer se u tom slučaju i šum pojačava i izjednačava po nivou sa govornim signalom. Kompromisno rešenje je da se umesto prenosne karakteristike 303, K_2 , primeni karakteristika 304, K_3 , kojom se smanjuje (komprimuje) ulazna dinamika prema zadatoj konstanti kompresije α . Karakteristika 304, K_3 se u decibelskom dijagramu opisuje linearnom funkcijom:

$$L_{out} = \alpha L_{in} + (1 - \alpha)L_{opt}, \quad (1)$$

gde su:

$$L_{in} = 10 \log(P_{in} / P_0),$$

$$L_{out} = 10 \log(P_{out} / P_0), \quad (2)$$

$$L_{opt} = 10 \log(P_{nom} / P_0),$$

i gde su: P_{in} – snaga signala na ulazu u kompresor, P_{out} – snaga signala na izlazu iz kompresora, P_{nom} – nominalna izlazna snaga na koju se želi podesiti izlazni signal a P_0 - referentna snaga u odnosu na koju se mere nivoi.

Kada je ulazni signal L_{opt} tada je nivo na izlazu takođe L_{opt} a linearno pojačanje A_{agc} za koje se ostvaruje ovaj proces treba da je 1. Ako je:

$$P_{out} = A_{agc}^2 P_{in}, \quad (3)$$

tada se za pojačanje A , ubacivanjem (2) i (3) u (1), dobija:

$$10 \log(A_{agc}^2 P_{in} / P_0) = 10\alpha \log(P_{in} / P_0) + (1 - \alpha) 10 \log(P_{nom} / P_0), \quad (4)$$

odnosno:

$$A_{agc} = \left(\frac{P_{nom}}{P_{in}} \right)^{(1-\alpha)/2}. \quad (5)$$

Iz (5) se vidi sledeća osobina pojačanja A_{agc} . Kada snaga ulaznog signala P_{in} teži nuli, uz uslov $0 < \alpha < 1$, tada pojačanje A_{agc} teži beskonačnosti. Ova osobina nije poželjna, jer u tom slučaju dolazi do velikog i nepotrebnog pojačanja vrlo slabih signala koji najčešće predstavljaju šum. Zbog toga se relacija (5) modifikuje dodavanjem člana δP_{nom} imeniocu te dobijamo:

$$A_{agc} = \left(\frac{P_{nom}}{P_{in} + \delta P_{nom}} \right)^{(1-\alpha)/2}. \quad (6)$$

Za ekstremno slabe signale, relacijom (6) se pojačanje ograničava na vrednost:

$$A_{agc \max} = \left(\frac{1}{\delta} \right)^{(1-\alpha)/2}. \quad (7)$$

Takođe, relacijom (6) prenosna karakteristika 304, K_3 , se modifikuje u karakteristiku 305, K_4 (slika 4). Na karakteristici K_4 uočavaju se tri karakteristične zone. Zona 1, kada je nivo ulaznog signala mnogo manji od vrednosti δP_{nom} . U tom slučaju je pojačanje A_{agc} konstantno prema relaciji (7). Očigledno je da ulazni signal ne treba da se nalazi u toj zoni, jer je tu pojačanje konstantno te nema delovanja AGC. Zona 2 predstavlja optimalnu radnu oblast. Za slabe signale pojačanje je limitirano na $A_{agc \max}$ dok se na delu oko L_{opt} primenjuje kompresija dinamike signala definisana konstantom

α . U zoni 3 je konstantan stepen kompresije ulaznog signala koji je definisan konstantom kompresije α .

Da bi ulazni signal doveli u jednu od zona 1, 2 ili 3, ulazni signal treba pojačati pre AGC bloka fiksnim pojačanjem, što karakteristiku 301, K_f , slika 3, pomera u položaj 306, K'_f .

Na slici 4 prikazan je blok dijagram AGC sistema. Njegov zadatak je: (1) da pojača slabe govorne signale a da oslabi previše jake signale prema unapred zadatoj karakteristici kompresije dinamike signala, (2) da na delovima ulaznog signala gde je prisutan samo eho signala, stacionaran šum ili konkurentni govornik-smetnja, smanji pojačanje kako bi se ove smetnje dovoljno utišale i (3) da utiša delove ulaznog signala gde su jednovremeno prisutni i koristan govorni signal i smetnje, a da pri tome očuva razumljivost govora.

Ulazni signal s_{in} , koji je jednak signalu s_{NR} iz bloka 205 sa slike 2, se množi u bloku 401 sa pojačanjem A_{agc} iz bloka 402, i tako se dobija izlazni signal $s_{out} = s_{AGC}$. Pojačanje A_{agc} je prema relaciji (6) određeno sa četiri parametra: P_{nom} - je fiksni parametar kojim se definiše željeni nivo izlaznog signala $s_{out} = s_{AGC}$, δ - je parametar koji ograničava pojačanje ulaznog signala $s_{in} = s_{NR}$ na vrednost $A_{agc\ max}$ (na primer: za $\delta = 0.001$ maksimalno moguće pojačanje je $A_{agc\ max} = 31.6$, jednačina (7)), P_{in} - snaga signala na ulazu u kompresor i α - konstanta kompresije koja definiše nagib karakteristike kompresije (slika 3). Relacija (6) se dalje modifikuje u smislu da je uvedena smena oznaka, pa je konstanta $\alpha = nagib$, i ona je složena funkcija vršne snage korisnog govornog signala \hat{P}_{dp} , i umesto P_{in} uvodi se estimacija snage ulaznog signala \tilde{P}_{in} , koja sadrži koristan signal aktuelnog govornika i sve signale smetnji. Ovom drugom smenom uvedena je funkcionalna zavisnost A_{agc} od signala smetnji, što se pokazalo veoma korisnim sa aspekta očuvanja kvaliteta izlaznog govornog signala. Sada relacija (6) postaje:

$$A_{agc} = \left(\frac{P_{nom}}{\tilde{P}_{in} + \delta P_{nom}} \right)^{0.5(1-nagib)} \quad (8)$$

Estimacija nagiba karakteristike kompresije *nagib* se vrši u bloku 403, dok se estimacija ulazne snage \tilde{P}_{in} vrši u bloku 404. Oba ova bloka zahtevaju estimaciju vršne

snage korisnog govornog signala \hat{P}_{dp} koja se određuje na sledeći način. Najpre se u bloku 405 određuje estimacija trenutne snage aktuelnog govornika iz odabranog pravca mikrofonskog niza, prema slici 2, na sledeći način:

$$\hat{P}_d = \sum_{f=f_{\min}}^{f_{\max}} \mathbf{C}'_f \mathbf{X}_f \mathbf{X}'_f \mathbf{C}_f = \sum_{f=f_{\min}}^{f_{\max}} |s_{in f}|^2, \quad (9)$$

gde su: $\mathbf{C}'_f = [c_{1,f} \ c_{2,f} \ \dots \ c_{N,f}]$ težinski koeficijenti superdirektivnog prostornog filtra SD-BF, $\mathbf{X} = [s_{AEC1} \ s_{AEC2} \ \dots \ s_{AECN}]'$, f_{\min} i f_{\max} su minimalna i maksimalna frekvencija u DFT spektru analiziranog govornog signala, respektivno. Zatim se u bloku 406 vrši procena vršne snage signala iz odabranog pravca $\hat{P}_{dp}(t)$ na sledeći način:

$$\hat{P}_{dp}(n) = \begin{cases} \hat{P}_d(n), & \text{ako je } \hat{P}_d(n) \geq \hat{P}_{dp}(n-1) \\ \alpha_d \hat{P}_{dp}(n-1) + (1 - \alpha_d) P_d(t), & \text{ako je } \hat{P}_d(n) < \hat{P}_{dp}(n-1) \end{cases} \quad (10)$$

gde je: n – aktuelni blok obrade signala i α_d - konstanta po vrednosti blizu 1.

U blok 404 kao informacija ulazi podatak, slika 4, o estimacija trenutne snage aktuelnog govornika \hat{P}_d (relacija (9)). Druga informacija jeste estimacija snage nepotisnutog eha \hat{P}_{eho} , koja se određuje na sledeći način:

$$\hat{P}_{eho} = \alpha_{eho} |\hat{y}|^2, \quad (11)$$

gde je: \hat{y} - estimirani signal eha dobijen u bloku 202, slika 2, dok je α_{eho} - konstanta potiskivanja eho signala.

Treća informacija koja ulazi u blok 404 jeste gruba procena snage difuznog šuma \hat{P}_n . Ova snaga se dobija obradom izlaznih signala s_{AEC1} do s_{AECN} iz bloka 202, blok AEC slika 2, i ulaznog signala $s_{in} = s_{NR}$ u blok AGC. Naime, signali s_{AEC1} do s_{AECN} ne sadrže akustički eho koji je potisnut u bloku 202, odnosno sadrže aktuelni govorni signal i sve ostale signale smetnji. Ako se od usrednjene snage signala s_{AEC1} do s_{AECN} oduzme estimirana snaga aktuelnog govornika \hat{P}_d ostaje gruba procena snage svih ostalih smetnji u snimanom ambijentu, pre svega difuznog šuma. Ovo se može izraziti sledećom relacijom:

$$\hat{P}_n = P - \hat{P}_d. \quad (12)$$

gde je P usrednjena snaga na izlazu iz bloka AEC i ona se određuje na sledeći način:

$$P = \frac{1}{N} \sum_{k=1}^N |s_{AECk}|^2. \quad (13)$$

Dakle, izlaz iz bloka 404 jeste estimacija snage \tilde{P}_n koja pored vršne snage govornog signala sadrži i estimacije snaga smetnji u ambijentu snimanja prema relaciji:

$$\tilde{P}_n = \hat{P}_{dp} + \hat{P}_n + \hat{P}_{eho}. \quad (14)$$

U bloku 403 vrši se estimacija nagiba karakteristike kompresije i izračunavanje veličine *nagib*. Postupak izračunavanja prikazan je na slici 5.

Analizom trajektorije vršne snage govornog signala izračunavaju se dva parametra: trend rasta, u bloku 501, i konveksnost trajektorije snage, u bloku 502, u cilju „meke odluke“ o prisustvu govora u ulaznom signalu, prema relacijama:

$$trend = 0.9375 \sum_{i=0}^3 (3/2 - i) \hat{P}_{dp}(n - i), \quad (15)$$

$$konveksnost_ = 1.5938 [\hat{P}_{dp}(n - 1) + \hat{P}_{dp}(n - 2) - \hat{P}_{dp}(n) - \hat{P}_{dp}(n - 3)], \quad (16)$$

$$konveksnost = \max\{konveksnost_ , 0\} \quad (17)$$

Zatim se izračunava složena promenljiva $\gamma(n)$, u bloku 503, sledećim postupkom:

$$\gamma_n = \frac{trend + konveksnost}{\hat{P}_{dp} + 0.006P_{nom}}, \quad (18)$$

$$\gamma(n) = \begin{cases} \gamma_n, & \text{ako je } \gamma_n > \gamma(n-1) \\ 0.93\gamma(n-1) + 0.07\gamma_n, & \text{ako je } \gamma_n \leq \gamma(n-1) \end{cases} \quad (19)$$

i veličina $\gamma_$, u bloku 504:

$$\gamma_ = \begin{cases} 1, & \text{za } \gamma(n) > 1 \\ \gamma(n), & \text{za } 0 < \gamma(n) < 1 \\ 0, & \text{za } \gamma(n) < 0 \end{cases} \quad (20)$$

Konačno, veličina *nagib*, koja predstavlja stepen kompresije dinamike signala i ulazi u relaciju (8), izračunava se u bloku 505 na osnovu prethodno izračunate veličine $\gamma_$ i veličine $nagib_{max}$ koja predstavlja zadatu maksimalnu vrednosti nagiba. Veličina *nagib* se računa relacijom

$$nagib = nagib_{max} \frac{1 + 2.5\gamma_^3}{nagib_{max} + 2.5\gamma_^3}. \quad (21)$$

Granični vrednosti veličine $nagib_$ su:

$$nagib = \begin{cases} nagib_{max} \frac{3.5}{nagib_{max} + 2.5} & \text{za } \gamma_ = 1, \text{ kompresija je velika kada } nagib \rightarrow 0 \\ 1 & \text{za } \gamma_ = 0, \text{ nema kompresije} \end{cases}$$

Time su određene sve veličine u jednačini (8) za izračunavanje pojačanja A_{agc} .

Sada se vidi uloga veličine \tilde{P}_m u izračunavanju A_{agc} , jednačina (8). Ukoliko je veći nivo snage zaostalog eho signala ili nivo difuznog ambijentalnog šuma, koji se ne može poništiti već samo redukovati, posredstvom AGC kontrole realizuje se manje pojačanje u odnosu na slučaj kada ne bi bilo prisutnih smetnji. Sa druge strane, ako u ulaznom signalu $s_{in} = s_{NR}$ nema govornog signala već samo rezidualnog šuma nakon redukcije šuma u NR bloku, blok 205 slika 2, tada \hat{P}_{ap} teži minimalnoj vrednosti kao i sve veličine u (15) do (21), što prouzrokuje smanjenje A_{agc} na minimalnu vrednost, odnosno $A_{agc} \rightarrow 1$.

U ovom pronalasku opisan je postupak obrade akustičkih i govornih signala u sistemu slobodne govorne komunikacije koji funkcioniše u punom dupleksu, a koji se odnosi na automatsku kontrolu izlaznog nivoa govornog signala. Specifičnost ovog pronalaska jeste njegova primena u sistemima na bazi mikrofonskih nizova, njegova potpuna realizacija u frekvencijskom domenu i mogućnost njegove primene i u drugim komunikacionim sistemima kao što su video-telefonski sistemi, telekonferencijski sistemi, spikerfoni u prostoriji ili kolima, komunikacija čovek-računar putem glasa, i td.

Postupci i tehnike obrade akustičkih i govornih signala u ovom pronalasku su nezavisni od broja mikrofona u mikrofonskom nizu, slika 2, i nalaze se pod kontrolom većeg broja parametara koji omogućavaju optimizaciju rešenja za različite aplikacije.

Postupci i tehnike obrade akustičkih i govornih signala u ovom pronalasku mogu se implementirati na različite načine. Na primer, ove tehnike mogu biti implementirane u hardveru, softveru ili kombinovano. U hardverskoj implementaciji mogu se koristiti specifična integrisana kola (ASIC), procesori za digitalnu obradu signala (DSP), programabilna logička kola (PLD ili FPGA) i druga elektronska kola projektovana tako da mogu izvršiti opisane funkcije u ovom pronalasku. U softverskoj implementaciji programski kodovi mogu biti memorisani u memorijskim jedinicama i izvršavani pomoću procesora kao što su PC, PDA, DSP, itd.

Na slici 6 prikazani su interfejsi AGC sistema kao nezavisnog hardversko/softverskog rešenja 600. Pored ulaza 601 i izlaza 602, za ulazni govorni signal i izlazni obrađeni govorni signal, ovo AGC rešenje ima i tri kontrolna ulaza 603, 604 i 605, preko kojih se u sistem uvode parametri $nagib_{max}$, P_{nom} i δ , respektivno, koji definišu optimalan rad AGC sistema. Pored toga, AGC sistem sadrži i $(N + 1)$ ulaza kao interfejs sa sistemom na bazi mikrofonskog niza od N mikrofona, namenjenog za

slobodnu („hands-free“) govornu komunikaciju u specifičnim telekonferencijskim uslovima primene.

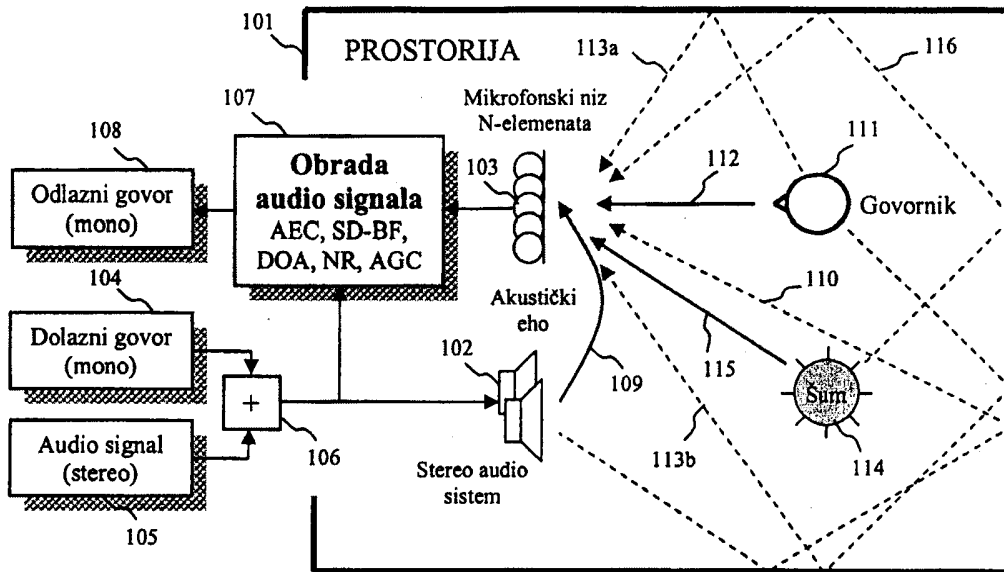
Detalji ovog pronalaska opisani ovde omogućavaju bilo kom stručnjaku u ovoj oblasti da generičke principe ovog pronalaska može implementirati u drugim sistemima za slobodnu govornu komunikaciju čime se ne izlazi iz okvira ovog pronalaska.

PATENTNI ZAHTEVI

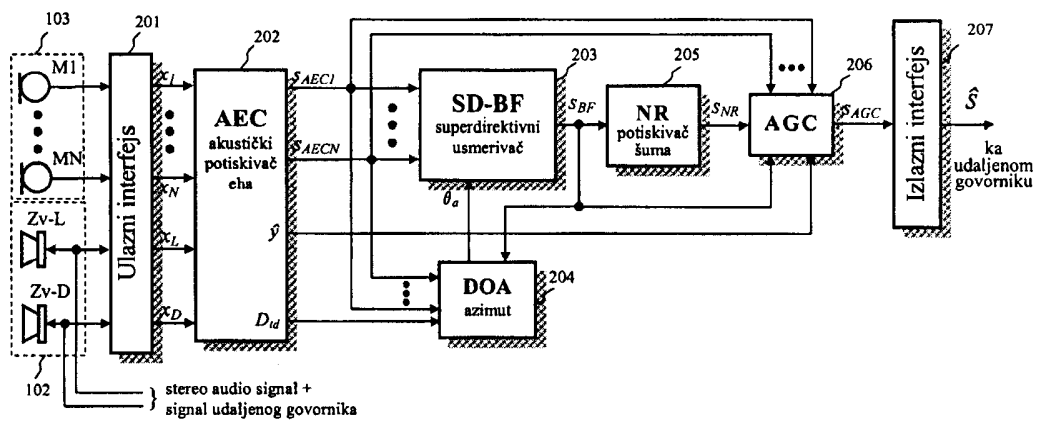
1. Postupak za automatsku regulaciju pojačanja (AGC) na osnovu očitavanja mikrofonskog niza **karakterisan time**, što sadrži:
adaptivno određivanje koeficijenta pojačanja, koji neposredno upravlja automatskom kontrolom pojačanja izlaznog signala i održava njegov nivo u definisanim granicama;
adaptivnu estimaciju nagiba karakteristike kompresije pomoću koje se komprimuje ulazni signal;
adaptivnu estimaciju vršne snage ulaznog signala, koja je osnovni ulazni parametar za estimaciju nagiba karakteristike kompresije i koeficijenta pojačanja;
adaptivnu estimaciju snaga signala smetnji na bazi očitavanja mikrofonskog niza u sistemu slobodne govorne komunikacije.
2. Postupak prema zahtevu 1 **karakterisan time**, što se kompletna obrada svih audio signala vrši u frekvencijskom domenu.
3. Postupak prema zahtevu 1 **karakterisan time**, što se koeficijent pojačanja ulaznog signala određuje adaptivno u zavisnosti od snage ulaznog govornog signala i snaga signala smetnji.
4. Postupak prema zahtevu 3 **karakterisan time**, što se koeficijent pojačanja ulaznog signala kontroliše pomoću dve konstante: nominalne izlazne snage na koju se želi podesiti izlazni signal i koeficijenta koji ograničava maksimalno pojačanje AGC sistema.
5. Postupak prema zahtevu 3 **karakterisan time**, što se koeficijent pojačanja ulaznog signala određuje na bazi adaptivne estimacije parametra nagiba karakteristike kompresije pomoću koje se komprimuje ulazni signal.
6. Postupak prema zahtevu 1 **karakterisan time**, što se koeficijent nagiba karakteristike kompresije estimira na bazi:
dve predefinisane konstante koje ograničavaju zonu delovanja AGC a čijim izborom se može optimizirati funkcija AGC u različitim aplikacijama, i
na bazi adaptivne estimacije vršne snage ulaznog signala koja neposredno određuje aktuelnu vrednost koeficijenta pojačanja.

7. Postupak prema zahtevu 6 **karakterisan time**, što se koeficijent nagiba karakteristike kompresije estimira na bazi dve predefinisane konstante: nominalne izlazne snage na koju se želi podesiti izlazni signal i koeficijenta koji definiše maksimalno dozvoljenu vrednost nagiba karakteristike kompresije.
8. Postupak prema zahtevu 6 **karakterisan time**, što se koeficijent nagiba karakteristike kompresije adaptivno estimira na bazi adaptivne estimacije vršne snage ulaznog signala i određivanja njene trajektorije u vremenu.
9. Postupak prema zahtevu 8 **karakterisan time**, što se koeficijent nagiba karakteristike kompresije određuje na bazi veličine trenda rasta trajektorije vršne snage ulaznog signala, adaptivno računatog na nekoliko prethodnih blokova analize vršne snage ulaznog signala.
10. Postupak prema zahtevu 8 **karakterisan time**, što se koeficijent nagiba karakteristike kompresije određuje na bazi veličine konveksnosti trajektorije vršne snage ulaznog signala, adaptivno računate na nekoliko prethodnih blokova analize vršne snage ulaznog signala.
11. Postupak prema zahtevu 8 **karakterisan time**, što se koeficijent nagiba karakteristike kompresije određuje na bazi „meke odluke“ o prisustvu govora u ulaznom signalu koristeći parametre trenda rasta i konveksnosti trajektorije vršne snage ulaznog signala.
12. Postupak prema zahtevu 1 **karakterisan time**, što se estimacija vršne snage ulaznog signala određuje adaptivno na takav način da se pri rastu snage usvaja trenutna estimacija snage, dok se pri opadanju snage koristi rekurzija sa određenom vremenskom konstantom.
13. Postupak prema zahtevu 1 **karakterisan time**, što se vrši estimacija snage nepotisnutog eha, na bazi estimacije snage eho signala u delu sistema slobodne govorne komunikacije koji vrši potiskivanje eha i primene procenjenog faktora potiskivanja eha do ulaza u AGC, a koja je potrebna za određivanje koeficijenta pojačanja AGC.

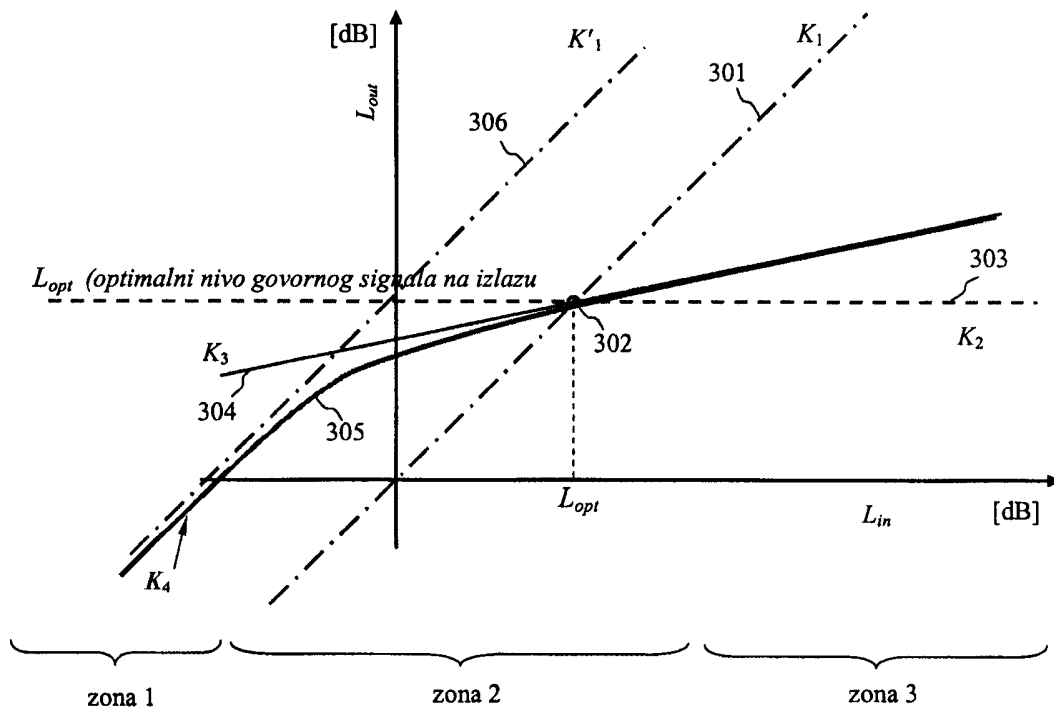
14. Postupak prema zahtevu 1 **karakterisan time**, što se vrši gruba procena snage difuznog šuma, u akustičkom ambijentu u kome mikrofonski niz snima, na bazi razlike estimirane srednje snage mikrofonskih signala, u kojima je izvršeno potiskivanje eha, i estimirane trenutne snage aktuelnog govornog signala, a koja je potrebna za određivanje koeficijenta pojačanja AGC.
15. Sistem za automatsku regulaciju pojačanja (AGC) na osnovu očitavanja mikrofonskog niza **karakterisan time**, što sadrži:
 - blok za adaptivno određivanje koeficijenta pojačanja, koji neposredno upravlja automatskom kontrolom pojačanja izlaznog signala i održava njegov nivo u definisanim granicama;
 - blok za adaptivnu estimaciju nagiba karakteristike kompresije pomoću koje se komprimuje ulazni signal;
 - blok za adaptivnu estimaciju vršne snage ulaznog signala, koja je osnovni ulazni parametar za estimaciju nagiba karakteristike kompresije i koeficijenta pojačanja;
 - blok za adaptivnu estimaciju snaga signala smetnji na bazi očitavanja mikrofonskog niza u sistemu slobodne govorne komunikacije.
16. Sistem prema zahtevu 15 **karakterisan time**, što je njegovo funkcionisanje optimizirano za primenu u sistemima slobodne govorne komunikacije na bazi mikrofonskih nizova u kojima je uključeno potiskivanje akustičkog eha (AEC) i lociranje govornika u prostoru na bazi usmerene karakteristike (BF) mikrofonskog niza, pri čemu broj mikrofona u nizu nije ograničavajući faktor.
17. Sistem prema zahtevu 15 **karakterisan time**, što sadrži jedan ulaz i jedan izlaz za aktuelni govorni signal i više ulaza na koje se dovode signali iz sistemima slobodne govorne komunikacije a koji omogućavaju estimaciju svih akustičkih smetnji u ambijentu snimanja mikrofonskog niza.
18. Sistem prema zahtevu 15 **karakterisan time**, što sadrži tri ulaza na kojima se postavljaju tri konstante koje definišu polje aktivnog funkcionisanja AGC, a to su: nominalna snaga izlaznog signala, dinamika promene pojačanja AGC sistema i maksimalni nagib karakteristike kompresije.



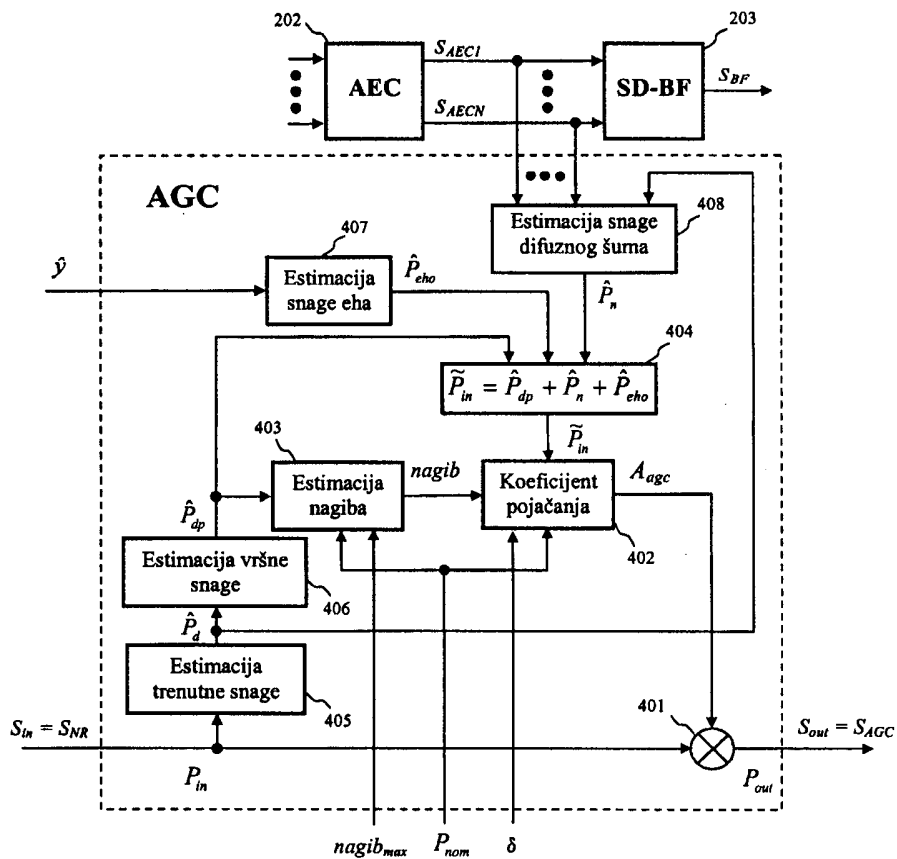
Slika 1.



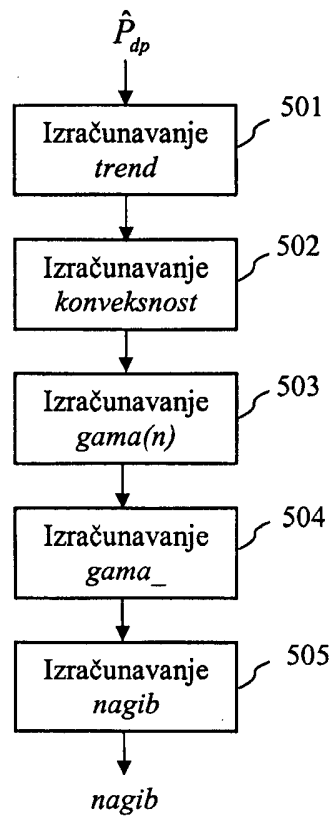
Slika 2.



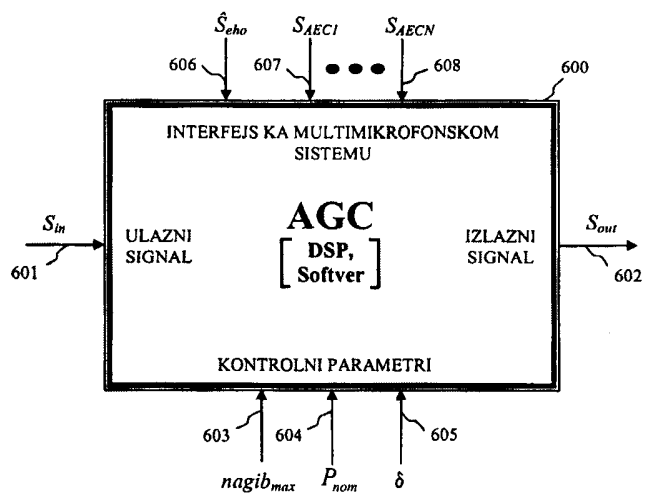
Slika 3.



Slika 4.



Slika 5.



Slika 6