

## **II Opis pronalaska**

### **Sistem za automatsko traženje zadatih audio sadržaja u radio i TV programima**

#### **Oblast tehnike na koju se pronalazak odnosi**

Pronalazak pripada oblasti IT (Informacione Tehnologije), odnosno informaciono-komunikacionih tehnologija, i predstavlja sistem koji, oslanjajući se na tehnike automatskog prepoznavanja govora, ima mogućnost brzog i jednostavnog praćenja komercijalnog sadržaja na velikom broju radio i televizijskih stanica (u daljem tekstu stanica).

#### **Tehnički problem**

U situaciji kada se na tržištu pojavljuje sve veći broj ponuđača proizvoda i usluga, markentiške kampanje sve više dobijaju na značaju. Sastavni deo svake markentiške kampanje je i adekvatna zastupljenost ponuđača u elektronskim medijima. Problem koji se javlja u toj situaciji je kako proveriti korektnost pri emitovanju reklamnog materijala u ugovorenim terminima. Proveravanje emitovanja u ugovorenim terminima do sada je zahtevalo veliko angažovanje ljudskih resursa.

Ozbiljne markentiške kampanje podrazumevaju i istraživanje tržišta i praćenje konkurencije. Praćenje medijske zastupljenosti konkurentskih proizvoda predstavlja još veći problem, jer ne postoji uvid u plan emitovanja.

Posebno je interesantno praćenje medijskog emitovanja autorski zaštićenog materijala, kako što su kompozicije, video klipovi i filmovi. Autorima, kao i udruženjima koja štite njihova prava, predstavlja velik problem praćenje emitovanja kompozicija na velikom broju radio i televizijskih stanica.

Tokom predizbornih kampanja postoji potreba da se prati zastupljenost političkih stranaka u medijima.

**Tehnički problem** se može definisati na sledeći način: kako omogućiti brzo i jednostavno praćenje medijskog sadržaja na velikom broju radio i televizijskih stanica uz minimalnu potrebu angažovanja ljudskih resursa. Podrazumeva se da se praćenje vrši iz prethodno snimljenog materijala. Kod televizijskog programa omogućeno je vizuelno praćenje, pa je postupak u izvesnoj meri ubrzan, dok je kod radija preslušavanje znatno sporije. U oba slučaja potrebno je znatno angažovanje ljudi, što dovodi u pitanje ekonomsku isplativost praćenja za “manje” oglašivače, kao i praćenje stanica sa nižom cenom reklamnog prostora.

**Predmet pronalaska** koji se štiti ovim patentom je novo tehničko rešenje za automatsko praćenje komercijalnog sadržaja na radio i televizijskim stanicama, kao jednostavan i pregledan pristup dobijenim rezultatima, što olakšava kontrolu, analizu i statističko praćenje određenih ponuđača proizvoda i usluga.

## **Stanje tehnike**

Razvoj tehničkih rešenja koja su omogućila relativno dostupnu masovnu memoriju velikog kapaciteta, zvučne kartice sa većim brojem ulaza i novih algoritama kompresije audio materijala, omogućilo je arhiviranje emitovanog programa velikog broja stanica. Međutim, za sam monitoring se trenutno koriste specijalizovana hardversko-softverska rešenja u čijoj je osnovi ipak čovek. Postoje posebni programi za reprodukciju snimljenog materijala kojima se upravlja preko nožnih papučica, pri čemu ruke ostaju slobodne za unos podataka. Na taj način se ubrzava preslušavanje snimljenog materijala i unos relevantnih podataka (reklama, muzičkih numera, izgovorenog teksta), ali ovo rešenje nije pogodno za praćenje velikog broja stanica, jer zahteva veliko angažovanje ljudskih resursa, a postoji i mogućnost greške ljudskog faktora usled zamora i pada koncentracije. Sve ovo poskupljuje praćenje i povećava verovatnoću pojave greške.

Razvoj informacionih tehnologija omogućio je povezivanje više računara u jedinstven sistem jednostavan za održavanje i nadogradnju. Delovi sistema se mogu nalaziti i na udaljenim lokacijama, što omogućava praćenje i lokalnih stanica.

Veliki napredak u oblasti automatskog prepoznavanja govora i doveo je do mogućnosti brzog i tačnog prepoznavanja željenog sadržaja u velikoj količini snimljenog materijala.

## **Izlaganje suštine pronalaska**

**Predmet pronalaska** koji se štiti ovim patentom je **Sistem za automatsko praćenje emitovanja reklamnog materijala** baziran na tehnikama prepoznavanja govora. Sistem predstavlja kompletno rešenje za snimanje, prepoznavanje, analizu rezultata i generisanje izveštaja.

**Suština pronalaska** je da se praćenje emitovanja komercijalnog sadržaja realizuje primenom tehnika iz oblasti prepoznavanja govora. Umesto čoveka koji preslušava snimljeni materijal i beleži pojavu određenog sadržaja, primenom ovih tehnika se audio zapis željenog sadržaja poredi sa snimljenim materijalom i u slučaju da dodje do poklapanja audio zapisa automatski se beleži pozicija gde se u materijalu nalazi prepoznati sadržaj.

## **Kratak opis slika nacрта**

**Predmet pronalaska** koji se štiti ovim patentom je **Sistem za automatsko praćenje emitovanja reklamnog materijala** baziran na tehnikama prepoznavanja govora. Sistem predstavlja kompletno rešenje za snimanje, prepoznavanje, analizu rezultata i generisanje izveštaja.

**Slika 1** – prikazuje šemu sistema za automatsko praćenje emitovanja reklamnog materijala

**Slika 2** – prikazuje predstavljanje zvučnog signala pomoću obeležja koja opisuju obvojnici spektra

**Slika 3** – prikazuje ilustraciju poređenja blokova

**Slika 4** – vizuelni prikaz rezultata prepoznavanja i dodavanje nove reklame na prepoznavanje

## Detaljan opis pronalaska

**Cilj pronalaska** je da se omogući brzo, jednostavno i tačno praćenje komercijalnih sadržaja na stanicama.

**Sistem** u osnovi pronalaska koji se štiti ovim patentom predstavlja novo tehničko rešenje za praćenje komercijalnog sadržaja na stanicama.

Jedan od načina ostvarivanja pronalaska prikazan je na Slici 1. Najpre antenski sistem prihvata TV/FM signal koji emituje predajnik (1), zatim se TV/FM signal prosleđuje TV/Radio tjunerima (2). Audio izlazi tjunera su povezani sa ulazima audio kartica računara SNIMANJE (3). Aplikacije za automatsko snimanje snimaju u fajlove zvučni signal sa ulaza audio kartica, a informacije o početku i trajanju emitovanog programa zabeleženog u fajlovima upisuju se u BAZU (4). Na radnoj stanici ADMINISTRIRANJE se vrši administriranje sistema koje podrazumeva izmene stanica čiji program se prati, kao i materijala koji se prepoznaje. Te informacije se, takođe, upisuju u BAZU (6). Aplikacije za automatsko prepoznavanje na računaru PREPOZNAVANJE dobijaju informacije iz BAZE (5) u kojim fajlovima se nalaze snimci emitovanja stanice koja se prati, za određeni period, kao i snimci materijala koji se prepoznaje. Rezultati prepoznavanja se, takođe, upisuju u BAZU (5).

Poređenje zvučnog signala omogućeno je poređenjem obeležja, koja na jedinstven način predstavljaju određenu zvučnu sekvencu. Najpre se signal подели na sekvence određene dužine obično manje od 30 ms (blokovi), jer su na tolikom periodu obeležja signala skoro stacionarna. Zatim se za svaku sekvencu izračunava određen broj obeležja, tako da je zvučni signal predstavljen nizom vektora obeležja. Za obeležja koja se koriste za reprezentaciju govornog signala najčešće se koriste **MFCC** (Mel Frequency Cepstral Coefficients) ili **LPC** (Linear Prediction Coefficients) koeficijenti. Ovi koeficijenti opisuju spektralnu obvojnici koja predstavlja glavnu

karakteristiku za razlikovanje signala i pri tome se redukuje ukupan broj parametara (u odnosu na slučaj kad bi se za poređenje koristili odbirci).

MFCC je baziran na poznatim varijacijama kritičnih frekvencija ljudskog uha. Filteri čija je uloga da izdvoje važne fonetske karakteristike govora, raspoređeni su linearno na nižim, a logaritamski na višim frekvencijama. Ovo je izraženo u mel-frekvencijskoj skali, sa linearnim frekvencijskim prostorom ispod 1000 Hz i logaritamskim iznad 1000 Hz. Prilikom formiranja blokova kontinualni zvučni signal se iseca na blokove dužina  $N$  odbirka, sa smicanjem u odnosu na prethodni blok za  $M$  ( $M < N$ ) odbiraka. Drugi blok počinje nakon  $M$  odbiraka prvog bloka i poklapa se sa njim u  $N - M$  odbiraka. Slično, treći blok počinje za  $2M$  odbiraka nakon prvog ( $M$  nakon drugog) i poklapa se sa njim na  $N - 2M$  odbiraka itd. Obično se uzima da je širina bloka (ili prozora) oko 30 ms ( $N$ ), dok je smicanje susednih blokova 10 ms ( $M$ ).

Sledeći korak je prozoriranje svakog zasebnog bloka, tako da se minimizuje diskontinuitet signala na početku i na kraju svakog zasebnog bloka. Ideja je da se minimizuje spektralno izobličenje, tako što se prozoriranjem signal postavlja na malu vrednost, na početku i na kraju svakog bloka. Ako se prozor definiše kao  $w(n)$ ,  $0 \leq n \leq N - 1$ , gde je  $N$  broj odbiraka u bloku, tada je rezultat prozoriranja signala

$$y_1(n) = x_1(n)w(n), \quad 0 \leq n \leq N - 1$$

Tipično se za predobradu signala koristi Hamingov prozor, koji ima formu:

$$w(n) = 0,54 - 0,46 \cos\left(\frac{2\pi n}{N - 1}\right), \quad 0 \leq n \leq N - 1$$

Svaki blok se konvertuje iz vremenskog u frekvencijski domen primenom Brze Furijeove transformacije (FFT – *Fast Fourier Transform*). FFT je brz algoritam za implementaciju diskretne Furijeove transformacije koja je definisana nad skupom  $N$  odbiraka  $\{x_n\}$ , kao:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j \frac{2\pi}{N} nk}, \quad k = 0, 1, 2, \dots, N-1$$

Ovde se  $j$  uzima kao imaginarna jedinica ( $j = \sqrt{-1}$ ). Generalno gledano  $X_k$  su kompleksni brojevi. Niz rezultata  $\{X_k\}$  se posmatra na sledeći način: nulta frekvencija odgovara vrednosti za  $k=0$ , pozitivne frekvencije  $0 < f < F_s/2$  odgovaraju vrednostima  $1 \leq k \leq N/2-1$ , dok negativne  $-F_s/2 < f < 0$  odgovaraju  $N/2+1 \leq k \leq N-1$ . Ovde  $F_s$  predstavlja frekvenciju uzorkovanja. Rezultat dobijen nakon ovog koraka se obično naziva spektar signala.

Mel-frekventna skala je linearni prostor za frekvencije ispod 1000 Hz, a logaritamski za frekvencije preko 1000 Hz. Za predstavljanje frekvencije na mel skali koristi se sledeća formula:

$$mel(f) = 2595 * \log_{10}(1 + f/700)$$

Jedan pristup simulacije subjektivnog spektra je korišćenje filtara, po jedan filter za svaku mel frekvenciju. Filtri su trougaoni propusnici opsega, a njihov prostorni raspored, kao i širina propusnog opsega su određeni konstantnim mel-frekvencijskim intervalom. Spektralna obvojnica od  $S(\omega)$  je zapravo snaga koja se dobija na izlazu ovih filtara kada je na njihovom ulazu  $S(\omega)$ .

U poslednjem koraku se konvertuje logaritam mel spektra u vremenski domen. To za rezultat daje mel koeficijente (MFCC). Spektralno predstavljanje zvučnog spektra signala obezbeđuje dobro predstavljanje lokalnih karakteristika spektra nad posmatranim blokom. Pošto su mel spektar koeficijenti (kao i njihovi logaritmi) realni brojevi, možemo ih konvertovati u vremenski domen korišćenjem diskretne kosinusne transformacije (DCT – *Discrete Cosine Transform*). Ako koeficijente dobijene u predhodnom koraku označimo sa  $\tilde{S}_k$ ,  $k = 1, 2, \dots, K$  možemo računati  $\tilde{c}_k$  (MFCC) kao:

$$\tilde{c}_n = \sum_{k=1}^K \left( \log \tilde{S}_k \right) \cos \left[ n \left( k - \frac{1}{2} \right) \frac{\pi}{K} \right], \quad n = 1, 2, \dots, K$$

Predstavljanje zvučnog signala pomoću obeležja koja opisuju obvojnici spektra prikazano je na Slici 2, gde je (1) zvučni signal u vremenskom domenu, (2) spektar tog signala i (3) predstavlja osam MEL (*Mel Frequency Cepstral*) koeficijenata koji opisuju obvojnici spektra. Pošto je zvučni signal podeljen na blokove od po ~30 ms, svaki blok je predstavljen sa 8 Mel koeficijenata.

Osnovna ideja LPC analize je pretpostavka da se trenutni uzorak zvuka može predvidjeti pomoću linearne kombinacije određenog broja prošlih uzoraka. Svi iznosi prošlih uzoraka ne doprinose jednako iznosu trenutnog uzorka, pa se moraju pre sumiranja pomnožiti odgovarajućim težinskim koeficijentima. Ti težinski koeficijenti su parametri linearnog prediktora. Cilj LPC analize je odrediti parametre prediktora. Koeficijenti prediktora ujedno su i koeficijenti IIR filtra koji modeluje spektralnu obvojnici zvučnog signala, odnosno frekventnu karakteristiku vokalnog trakta. Amplitudno-frekventna karakteristika takvog filtra predstavlja obvojnici vremenski kratkotrajnog spektra signala za određeni okvir analize.

Analiza govornog signala, podrazumeva izračunavanje težinskih parametara linearne predikcije direktno iz uzorka govornog signala. Svaki uzorak govornog signala može se aproksimirati linearnom kombinacijom prethodnih uzoraka:

$$\hat{s}(n) = \sum_{k=1}^p a_k s(n-k)$$

Greška predikcije  $e(n)$  definisana je kao:

$$e(n) = s(n) - \hat{s}(n) = s(n) - \sum_{k=1}^p a_k s(n-k)$$

Zbog vremenski promenljive prirode govornog signala, koeficijenti prediktora moraju biti određeni iz kratkog segmenta govornog signala unutar kojeg su spektralna svojstva signala stalna. Taj segment je označen sa  $s_n(m)$ , a radi se o delu signala  $s(n)$  u okolini indeksa  $n$ , gde je ta okolina određena rasponom indeksa  $m$ , tj. važi:

$$s_n(m) = s(m+n)$$

Osnovna ideja linearne predikcije je pronalaženje takvih koeficijenata prediktora koji će minimizovati srednju kvadratnu grešku predikcije na tom posmatranom segmentu govornog signala.

$$E_n = \sum_m e_n^2(m) = \sum_m (s_n(m) - \bar{s}(m))^2 = \sum_m \left[ s_n(m) - \sum_{k=1}^p a_k s_n(m-k) \right]^2$$

Vrednosti parametara  $a_k$  koji minimiziraju  $E_n$  prema gornjoj relaciji mogu se odrediti tako da se parcijalne diferencijalne jednačine  $E_n$  po nepoznatim koeficijentima izjednače s nulom, tj.

$$\frac{\partial E}{\partial a_k} = 0 \text{ za } k = 1, 2, \dots, p$$

Iz ovog uslova sledi sistem od  $p$  linearnih jednačina po nepoznatima  $a_1$  do  $a_p$ . Rešavanjem ovog sistema jednačina dobija se skup optimalnih koeficijenata prediktora  $a_k$  koji minimizuju srednju kvadratnu grešku predikcije  $E_n$  za segment  $s_n(m)$ .

Da bi se izbegla zavisnost rezultata poređenja od intenziteta signala, preporučljivo je da se za poređenje koriste delta (prvi izvod u vremenu) i delta-delta (drugi izvod) vrednosti obeležja. Ovi delta koeficijenti, ne samo da smanjuju zavisnost obeležja od intenziteta signala, već i od mnogih drugih linearnih i relativno stacionarnih izobličenja, kao što su: stacionarni šum, frekvencijska izobličenja po različitim kanalima i sl. Izvodi se računaju preko regresionih koeficijenata. Regresioni koeficijenti dobijaju se kada se funkcija promene određenog obeležja u vremenu aproksimira nekim polinomom po kriterijumu minimalne kvadratne greške, a zatim se nađe vrednost prvog izvoda te funkcije u željenoj tački. Ovo se radi u prozoru određene širine. Kada se aproksimacija vrši polinomom drugog reda dobije se:

$$r_{i,j} = \frac{\sum_{q=-Q}^Q q \cdot P_{i+q,j}}{\sum_{q=-Q}^Q q^2}$$



gde je  $2Q+1$  širina prozora (u blokovima), a  $p_{k+q,j}$  je vrednost obeležja,  $i$  je broj bloka, a  $j$  indeks obeležja.

Poređenje zvučnog zapisa reklame sa zvučnim zapisom snimljenog materijala zasniva se na izračunavanju niza euklidskih rastojanja između vektora obeležja reklame i adekvatnih vektora iz snimljenog materijala. Suma tog niza predstavlja rastojanje reklame od zvučnog zapisa snimka emitovanja. Ukoliko je izračunato rastojanje ispod određene vrednosti (praga) smatra se da snimak sadrži emitovanje te reklame. Na osnovu pozicije u snimku gde je reklama pronađena izračunava se vreme emitovanja reklame. Kako se radi o euklidskom rastojanju, potrebno je utvrditi težinske koeficijente za pojedina obeležja, kako bi se njihov uticaj ujednačio. Ovo je urađeno eksperimentalno.

Ilustracija ovog poređenja prikazana je na Slici 3, gde je (1) niz  $n$  blokova obeležja koji predstavljaju reklamu  $r$  koja se prepoznaje, dok je (2) niz blokova obeležja koji predstavljaju snimak  $s$  emitovanja u kojem se reklama prepoznaje. Svaki blok u reklamama  $r_i$  ( $i=1, \dots, n$ ), sadrži vektor delta i delta-delta obeležja koji reprezentuju ~30 ms zvučnog snimka. Na Slici 3 je prikazano poređenje koje je započeto u bloku snimaka  $s_m$ . Svaki vektor  $r_i$  bloka poredi se sa adekvatnim vektorom  $s_j$  bloka izračunavanjem euklidskog rastojanja. Ukoliko srednje rastojanje poređenja ne zadovoljava kriterijum prepoznavanja, postupak se ponavlja sa početkom u bloku  $s_{m+1}$  itd. U slučaju da je kriterijum prepoznavanja zadovoljen, smatra se da je reklama prepoznata u snimku sa početkom u  $s_m$  bloku.

Poređenjem različitih reklama sa emitovanim signalom utvrdilo se da srednja vrednost euklidskog rastojanja jako varira od reklame do reklame. Ovo važi kako za delove signala gde nije prisutna reklama koja se traži, tako i za onaj deo u kojem je reklama emitovana. Zbog toga nije moguće usvojiti jedinstveni prag za sve reklame koje se traže u snimljenom materijalu. Tako da se prag na osnovu kojeg se procenjuje da li je određena reklama emitovana u datom momentu određuje dinamički. Izračunava se i rastojanje reklame od snimka za niz blokova (npr. 50) ispred i iza trenutne pozicije. Prag predstavlja jednu polovinu od srednje vrednosti tih rastojanja.

Ovaj način poređenja može biti prilično spor u slučajevima kada se traže reklame ili muzičke numere dužine trajanja od po nekoliko minuta. Iz tog razloga je

poželjno optimizovati brzinu prepoznavanja. Prepoznavnje se može ubrzati korišćenjem dva nivoa poređenja. U prvom nivou se koristi manji broj blokova (npr. 50) za izračunavanje i povećava se prag, tako da se brže odbacuju delovi snimljenog materijala sa velikom distancom. Ukoliko u određenom bloku snimka distanca zadovoljava kriterijum prvog nivoa, prelazi se na drugi nivo gde se porede svi blokovi reklame sa odgovarajućim uzastopnim blokovima u snimku, na ranije opisan način. Reklama je pronađena u snimku samo ako je zadovoljen kriterijum u drugom nivou prepoznavanja.

Da bi se postigli što bolji rezultati potrebno je napraviti izbor blokova u reklami koji će se porediti u prvom nivou. Ideja je da to budu najdistinktivniji blokovi, odnosno oni koji imaju najveće module vektora obeležja. Jedan od načina da se oni odaberu je poređenje blokova (vektora obeležja) reklame sa nula vektorima, a zatim se izabere  $N$  blokova koji imaju najveću distancu. U prvom nivou se tih  $N$  blokova poredi sa odgovarajućim blokovima u snimljenom materijalu.

Na primer, ukoliko snimak reklame ima 500 blokova (~ 15 sekundi) prvo se na opisan način izabere 50 blokova  $r_5, r_{35}, r_{41}, \dots$ , a ta reklama se traži u snimku dužine 120000 blokova (~ 1 sat) i reklama se traži u 40000. bloku (oko 20-og minuta), tada se distanca računa poređenjem izabranih blokova reklame sa blokovima iz snimka  $s_{40004}, s_{40034}, s_{40040} \dots$

Ukoliko je prag prvog nivoa zadovoljen, prelazi se na drugi nivo gde se porede svi blokovi reklame  $r_1, r_2, \dots, r_{500}$  sa blokovima snimljenog materijala  $s_{40000}, s_{40001}, \dots, s_{40499}$ . Ako izračunata distanca zadovoljava kriterijum drugog nivoa, smatra se da je početak reklame nalazi u 40000. bloku snimka (20 minuta od početka snimka). Na osnovu podatka o datumu i vremenu vezanih za početak snimka emitovanja, npr. 07. jun 2006. u 14:00:00 možemo da zaključimo da je data reklama emitovana 07. juna 2006. u 14:20:00.

Najbolji rezultati gore opisanog prepoznavanja mogu se postići u okviru većeg sistema koji u sebi sadrži deo za snimanje, bazu podataka kao i korisničku aplikaciju koja omogućava lako upravljanje i održavanje sistema.

Deo za snimanje emitovanog programa radio i televizijskih stanica sastoji se od FM/TV tjunera i aplikacija za snimanje. Aplikacije snimaju audio fajlove i u bazu podataka upisuju vreme i dužinu snimljenog programa za stanice koje su na praćenju.

Korisnička aplikacija omogućava definisanje stanica čiji se program prati, kao i reklame koje se tom prilikom prepoznaju. Omogućava jednostavan pregled rezultata prepoznavanja. Zahvaljujući podacima koji se čuvaju u bazi moguće je lako generisanje izveštaja koji odgovaraju zahtevima korisnika.

Aplikacije za automatsko prepoznavanje na osnovu zadatih perioda vrše prepoznavanje reklama u snimcima emitovanja određenih stanica, a rezultate upisuju u bazu.

Ovaj koncept omogućava laku proširivost sistema dodavanjem novih stanica na praćenje i dodatnih računara na kojima će se izvršavati aplikacije za automatsko prepoznavanje. Preduslov za dodavanje novih stanica je proširenje hardverskih kapaciteta sistema – dodavanje FM/TV tjunera i audio kartica na čije ulaze će tjuneri biti vezani. U slučaju da dodatni broj stanica prevazilazi hardverska ograničenja računara u sistemu (broj PCI slotova), sistem se može proširiti proizvoljnim brojem dodatnih računara. Nakon toga se u korisničkoj aplikaciji definišu nove stanice i eventualno novi računari, kao i materijal koji treba da se prati na tim stanicama. Sa dodavanjem novih stanica na praćenje povećava se i vreme potrebno za dobijanje kompletnih rezultata pretrage svih stanica u sistemu. Ukoliko je potrebno, to se vreme može skratiti dodavanjem novih računara koji će se koristiti za pretragu. Registracija ovih računara u sistem nije potrebna. Dovoljno je, na tim računarima, instalirati i konfigurisati aplikacije za prepoznavanje i one se automatski javljaju bazi odakle dobijaju snimke emitovanja stanica i materijala koji treba prepoznavati, a rezultate prepoznavanja beleže u bazu.

Pošto postoji zakonska obaveza da se reklame emituju u okviru reklamnih blokova, u slučaju praćenja kompletnog reklamnog sadržaja na određenoj stanici, na osnovu rezultata prepoznavanja je moguće automatsko detektovanje pojave „rupa“ u reklamnim blokovima, što ukazuje na pojavu nove reklame na toj stanici. U klijentskoj aplikaciji, gde se uočava ova pojava, moguće se jednostavno izdvajanje snimka reklame iz emitovanog sadržaja i dodavanje te reklame na praćenje. Ovo je prikazano na slici 4.

## **Način industrijske i druge primene pronalaska**

Osnovna oblast primene sistema koji je predmet pronalaska koji se štiti ovim patentom je praćenje medijskog prostora na radio i televizijskim stanicama.

Sistem je u najvećoj meri pogodan za korišćenje u medija monitoring agencijama, koje će prikupljene informacije plasirati svojim klijentima. Pošto je sistem potpuno automatizovan, jedina funkcija koja zahteva ljudsko angažovanje je ubacivanje novih sadržaja na pretragu i dodavanja novih radio i televizijskih stanica u sistem. Ovo omogućuje i praćenje stanica za koje nije postojala ekonomska isplativost dok je za to bilo potrebno angažovanje velikog broja ljudi.

Klijentima koji su zakupili reklamni prostor sada će biti dostupne potpune informacije o poštovanju dogovorenih termina emitovanja. Takođe su dostupne informacije i o medijskoj zastupljenosti konkurentskih proizvoda, što može biti značajno za osmišljanje ili izmene sopstvenih reklamnih kampanja.

Ovaj sistem u velikoj meri olakšava vlasnicima autorskih prava da prate emitovanje sadržaja na osnovu kojih im se uplaćuju prihodi.

Mogućnost distribuirane organizacije sistema i daljinsko upravljanje, olakšava praćenje i lokalnih stanica.

Primenom sistema koji je predmet pronalaska koji se štiti ovim patentom se na lakši, brži i pouzdaniji način prati medijski prostor na stanicama, što utiče i na nižu cenu i veću upotrebnu vrednost dobijenih informacija.

Potpis podnosioca prijave

---

### III Patentni zahtev

1. Sistem za automatsko traženje zadatih audio sadržaja u radio i TV programima, *naznačen time*, što snima audio signale emitovanih radio i TV programa i poredi sadržaje niza kratkih vremenskih segmenata iz zadatih i snimljenih audio signala poređenjem odgovarajućih vektora spektralnih obeležja i rezultate pretrage beleži u bazu podataka.

2. Sistem za automatsko traženje zadatih audio sadržaja prema zahtevu 1, *naznačen time*, što ljudsku asistenciju svodi na zadavanje audio signala koji treba da se pronade i snimljenog audio signala u okviru kojeg treba da se vrši pretraga.

3. Sistem za automatsko traženje zadatih audio sadržaja prema zahtevu 1, *naznačen time*, što može simultano da vrši veći broj zadataka audio pretrage i snimanja radio i TV signala u širem geografskom području na distribuiranoj arhitekturi i sa daljinskim upravljanjem.

Potpis podnosioca prijave

---

## IV Apstrakt

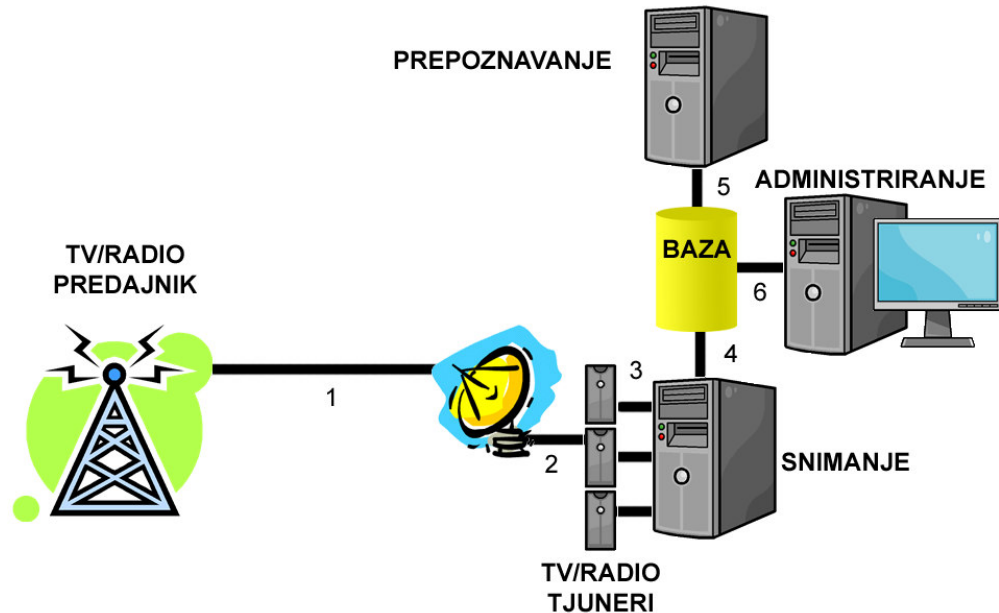
**Sistem za automatsko praćenje emitovanja reklamnog materijala baziran na tehnikama prepoznavanja govora** predstavlja kompletno rešenje za snimanje, prepoznavanje, analizu rezultata i generisanje izveštaja.

Prednost ovog sistema je što se kompleksni postupak praćenja emitovanih sadržaja koji zahteva velike ljudske resurse, zamenjuje sa nekoliko jednostavnih postupaka potrebnih za administriranje sistema. Umesto čoveka koji preslušava snimljeni materijal i beleži pojavu određenog sadržaja, primenom ovih tehnika se audio zapis željenog sadržaja automatski poredi sa snimljenim materijalom i u slučaju da dođe do poklapanja audio zapisa beleži se pozicija gde se u materijalu nalazi prepoznati sadržaj.

Potpis podnosioca prijave

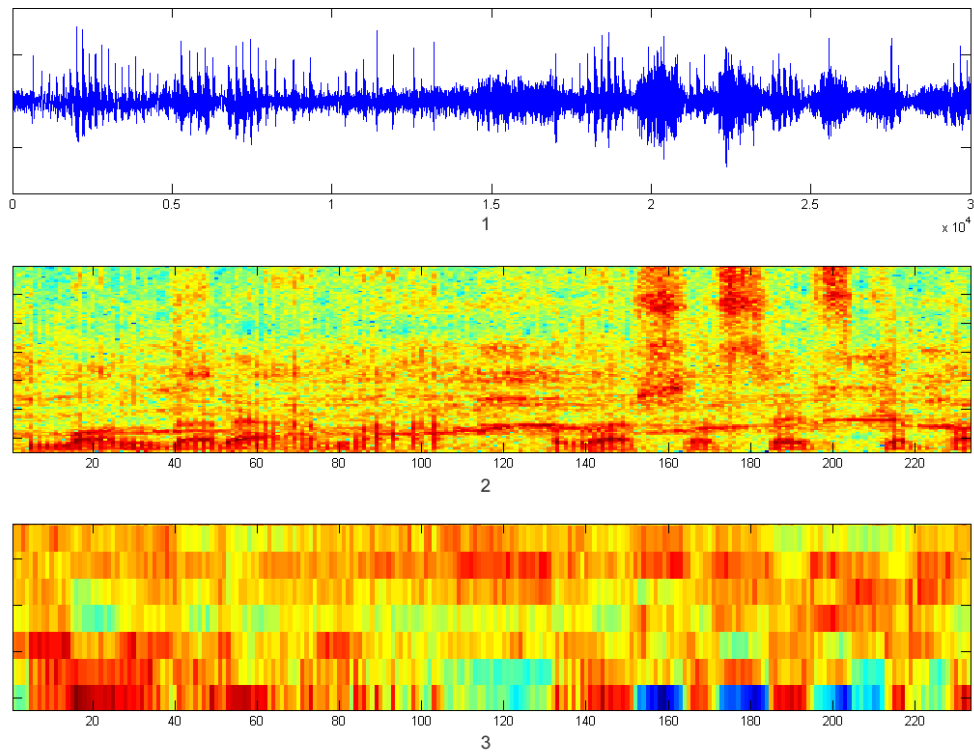
---

## Nacrt pronalaska



**Slika 1** – Sistem za automatsko praćenje emitovanja reklamnog materijala

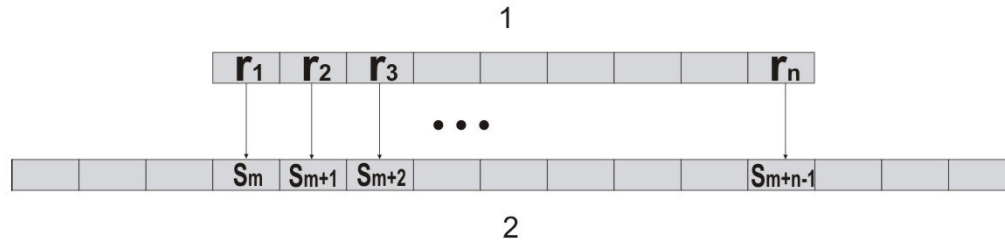
- (1) Antenski sistem prihvata TV/Radio signal koji emituje predajnik.
- (2) TV/Radio tjuneri dobijaju signal sa antenskog sistema.
- (3) Aplikacije za automatsko snimanje snimaju u fajlove zvučni signal sa ulaza audio kartica koji dobijaju sa tjunera.
- (4) Informacije o početku i trajanju emitovanog programa zabeleženog u fajlovima upisuju se u bazu podataka
- (5) Aplikacije za automatsko prepoznavanje dobijaju informacije iz baze podataka u kojim fajlovima se nalaze snimci emitovanog programa stanice koja se prati, za određeni period, kao i snimci materijala koji se prepoznaje. Rezultati prepoznavanja se, takođe, upisuju u bazu podataka.
- (6) Administriranje sistema podrazumeva izmene stanica koje se prate, kao i materijala koji se prepoznaje. Te informacije se, takođe, nalaze u bazi podataka.



**Slika 2** – Predstavljanje zvučnog signala pomoću obeležja koja opisuju obvojnici spektra

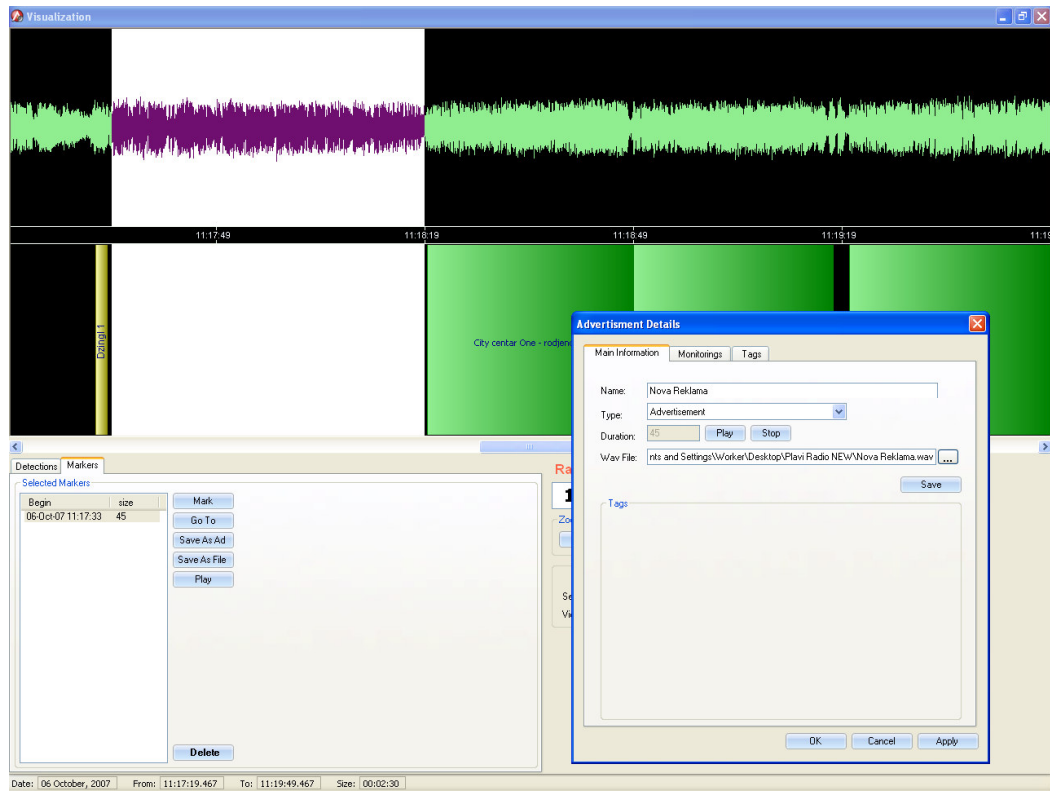
- (1) Zvučni signal u vremenskom domenu.
- (2) Spektar signala sa slike (1).
- (3) Osam **MFCC** (*Mel Frequency Cepstral*) koeficijenata koji opisuju obvojnici spektra. Pošto je zvučni signal podeljen na blokove od po  $\sim 30$  ms, svaki blok je predstavljen sa 8 Mel koeficijenata, čiji se prvi i drugi izvod u vremenu koristi za prepoznavanje.





**Slika 3** – Ilustracija poređenja blokova

- (1) Niz  $n$  blokova obeležja koji predstavljaju reklamu koja se prepoznaje. Svaki  $r$  blok sadrži vektor delta i delta-delta obeležja koji reprezentuju ~30 ms zvučnog snimka.
- (2) Niz blokova obeležja koji predstavljaju snimak emitovanja u kojem se reklama prepoznaje. Na slici je prikazano poređenje koje je započeto u bloku snimaka  $s_m$ . Svaki vektor  $r$  bloka poredi se sa adekvatnim vektorom  $s$  bloka izračunavanjem euklidskog rastojanja. U slučaju da je kriterijum prepoznavanja zadovoljen smatra se da je reklama prepoznata u snimku sa početkom u  $s_m$  bloku. Ukoliko srednje rastojanje poređenja ne zadovoljava kriterijum prepoznavanja, postupak se ponavlja sa početkom u bloku  $s_{m+1}$  itd.



**Slika 4** – Vizuelni prikaz rezultata prepoznavanja i dodavanje nove reklame na prepoznavanje

Potpis podnosioca prijave

---