

DIGITALNI WATERMARKING GOVORNIH SIGNALA ZASNOVAN NA VREMENSKI-PROMJENLJIVOM FILTRIRANJU

A DIGITAL WATERMARKING OF SPEECH SIGNAL BASED ON TIME-VARYING FILTERING

Nikola Žarić, Irena Orović, Srdjan Stanković, Elektrotehnički fakultet Podgorica
irenao, zaric, srdjan@cg.ac.yu
Cornel Ioana, ENSIETA, Brest, France, ioanaco@ensieta.fr

Sadržaj – U radu je predložen novi pristup u kreiranju i umetanju digitalnog watermark-a u govorni signal. Oblast pogodna za umetanje watermark-a izdvojena je korišćenjem vremensko-frekvencijske reprezentacije signala. Koncept vremenski-promjenljivog filtriranja iskorišćen je za vraćanje signala iz vremensko-frekvencijskog u vremenski domen. Umetanje i detekcija dobijenog watermark signala izvršena je u vremenskom domenu, i ilustrovana je na primjeru.

Abstract – A new approach in digital watermarking of speech signal is proposed. The suitable region for watermark embedding is found out by using the time-frequency representation. The watermark is created to correspond with the specified region in the time-frequency plane. The concept of the time-varying filtering is used in order to retrieve back the signal from the time-frequency domain. The procedure of watermark embedding, as well as watermark detection is performed in the time domain.

Ključne riječi: digitalni watermark, vremensko-frekvencijska analiza, vremenski promjenljivo filtriranje
Keywords: digital watermarking, time-frequency analysis, time-varying filtering

1. UVOD

Razvoj digitalnih medija izazvao je potrebu za kreiranjem metoda kojima bi se zaštitila autorska prava vlasnika. Konvencionalni kriptografski sistemi omogućavaju jedino vlasniku ključa pristup šifrovanim podacima. Međutim, ukoliko se ti podaci dešifruju gubi se mogućnost praćenja njihove reprodukcije i retransmisije. Da bi se prevazišli nedostaci konvencionalne kriptografije, uvedena je tehnika digitalnog watermarking-a. Princip watermarking-a zasniva se na umetanju digitalnog koda (watermark-a) u multimedijalni sadržaj (audio signal, sliku, video signal), kao i korišćenje istog za dokazivanje vlasništva. Umetanje digitalnog watermark-a vrši se na način koji ne rezultira perceptualnom modifikacijom digitalnih podataka. Istovremeno, watermark mora da bude robustan na standardne obrade podataka, koje ne smanjuju kvalitet obrađivanog materijala [1]. S toga, procedura umetanja watermarka treba da obezbijedi zadovoljenje ova dva protivurječna zahtjeva.

Kada se govori o watermarking-u govornog signala, potrebno je napomenuti da postoji čitav niz naprednih tehnika. Generalno govoreći, watermark može da bude umetnut u vremenskom ili nekom od transformacionih domena. Jedna od jako zastupljenih tehnika watermarking-a govornog signala je spread-spectrum tehnika [1], [2]. Ona podrazumijeva dodavanje pseudo-slučajne sekvence originalnom signalu, dok se detekcija tako umetnutog watermark-a vrši korišćenjem korelacione funkcije. U drugu kategoriju spada čitav niz samo-markirajućih tehnika, a kao

jedan specifičan slučaj razvijen je metod vremenskog skaliranja [4].

U ovom radu dat je nešto drugačiji pristup watermarking-u govornog signala. Naime, da bi se lokalizovao region pogodan za umetanje watermark-a, korišćen je spektrogram kao vremensko-frekvencijska reprezentacija govornog signala. U nastavku, korišćenjem tehnike vremenski-promjenljivog filtriranja omogućeno je kreiranje efikasnog watermark signala koji se dodaje govornom signalu, u vremenskom domenu. Detekcija umetnutog watermark-a realizovana je pomoću korelacione funkcije u vremenskom domenu.

2. TEORIJSKA OSNOVA

U narednom dijelu dat je kratak osvrt na vremensko-frekvencijske distribucije koje su korišćene pri realizaciji predložene watermarking šeme. Takođe, dato je objašnjenje tehnike vremenski-promjenljivog filtriranja zasnovanog na vremensko-frekvencijskoj distribuciji.

Vremensko-frekvencijska reprezentacija omogućava istovremenu analizu signala i u vremenu i po frekvencijama. Osnova svih vremensko-frekvencijskih reprezentacija je Short Time Fourier-ova Transformacija (STFT), i definiše se na sljedeći način:

$$STFT(t, \omega) = \int_{-\infty}^{\infty} x(t + \tau)w(\tau)e^{-j\omega\tau} d\tau, \quad (1)$$

gdje $w(\tau)$ predstavlja funkciju prozora širine τ .

Diskretna forma relacije (1) zapisuje se kao:

$$STFT(n, k) = \sum_{m=-N/2}^{N/2} w(m)x(n+m)e^{-j2\pi mk/N}, \quad (2)$$

gdje su n i k diskretne vrijednosti vremena i frekvencije, respektivno, dok je $w(m)$ funkcija prozora u diskretnom obliku.

STFT je veoma jednostavna za realizaciju. Naime, za određeni vremenski trenutak STFT se računa korišćenjem algoritma brze Fourier-ove transformacije (Fast Fourier Transform-FFT) [5]:

$$STFT(n, k) = DFT\{w(m)x(n+m)\}, \quad (3)$$

gdje je sa DFT označena diskretna Fourier-ova transformacija.

Energetska verzija STFT predstavlja spektrogram koji se definiše kao:

$$SPEC(n, k) = |STFT(n, k)|^2. \quad (4)$$

Za multikomponentne signale kakav je govorni signal, STFT se može predstaviti kao suma STFT svih komponenti [5]:

$$STFT(n, k) = \sum_{m=1}^M STFT_m(n, k), \quad (5)$$

gdje je M ukupan broj komponenti signala. Shodno tome, ukoliko se komponente signala ne preklapaju u vremensko-frekvencijskom domenu, spektrogram multikomponentnog signala se može izračunati kao:

$$SPEC(n, k) = \sum_{m=1}^M SPEC_m(n, k). \quad (6)$$

Imajući u vidu da ne postoji mogućnost direktne rekonstrukcije vremenskog oblika signala iz vremensko-frekvencijskog domena, koristi se tehnika vremenski-promjenljivog filtriranja. Ova tehnika je, originalno, uvedena za filtriranje zašumljenih signala [6], i definisana je kao:

$$Hx(t) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2}, t - \frac{\tau}{2})x(t + \tau)d\tau, \quad (7)$$

gdje je $x(t)$ signal, a $h(t+\tau/2, t-\tau/2)$ impulsni odziv vremenski-promjenljivog filtra.

Generalizovana transfer funkcija vremenski-promjenljivog filtra poznata kao Weyl simbol [7], definiše se na osnovu impulsnog odziva kao:

$$L(t, \omega) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2}, t - \frac{\tau}{2})e^{-j\omega\tau}d\tau. \quad (8)$$

Na osnovu relacija (7) i (8), i korišćenjem Parsevalove teoreme, signal na izlazu vremenski-promjenljivog filtra može da se rekonstruiše korišćenjem relacije:

$$Hx(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} L_H(t, \omega)STFT(t, \omega)d\omega \quad (9)$$

3. VREMENSKO-FREKVENCIJSKI ZASNOVANA WATERMARKING TEHNIKA

Polazeći od spektograma kao vremensko-frekvencijske reprezentacije pogodne za analizu govornih signala, i pretpostavljajući da su sve komponente spektograma signala skoncentrisane u regionu D^2 vremensko-frekvencijskog domena, funkcija L_H može se zapisati kao:

$$L_H(t, \omega) = \begin{cases} 1, & \text{za } (t, \omega) \in D^2 \\ 0, & \text{za } (t, \omega) \notin D^2 \end{cases} \quad (10)$$

Ovako definisana, funkcija L_H pruža potpunu informaciju o pozicijama svih komponenti signala u vremensko-frekvencijskoj ravni. Priroda govornog signala je takva da se pri umetanju watermark-a, zbog perceptualne osjetljivosti izbjegavaju visoko-frekventne komponente, kao i pauze u govornom signalu. S obzirom na to, potrebno je definisati adekvatnu podoblast D_M oblasti D^2 , koja će biti podesna za watermarking. Ukoliko se podoblast D_M definiše tako da obuhvata samo najjače komponente u vremensko-frekvencijskom domenu, D_M se može zapisati na sljedeći način:

$$D_M = \begin{cases} 1, & \text{za } SPEC_s(t, \omega) > \xi \\ 0, & \text{za } SPEC_s(t, \omega) \leq \xi \end{cases} \quad (11)$$

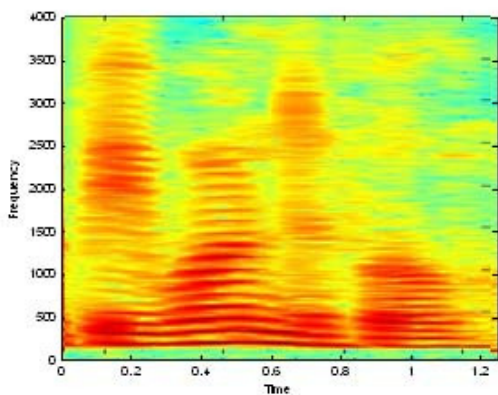
gdje je sa $SPEC_s(t, \omega)$ označen spektrogram govornog signala, a ξ predstavlja prag koji se definiše kao:

$$\xi \geq \overline{SPEC_s}(t, \omega). \quad (12)$$

Modifikovana transfer funkcija L_H definisana u oblasti D_M se može zapisati kao:

$$L_M(t, \omega) = \begin{cases} 1, & \text{za } (t, \omega) \in D_M \\ 0, & \text{za } (t, \omega) \notin D_M \end{cases} \quad (13)$$

Spektrogram govornog signala, na osnovu kojeg se definiše oblast pogodna za dobijanje funkcije L_M , prikazan je na Slici. 1.



SI.1. Spektrogram govornog signala

U nastavku, funkcija L_M biće korišćena za dobijanje watermark-a. Polazeći od pseudo-slučajne sekvence p , njena filtrirana forma koja će služiti kao watermark, dobija se korišćenjem formule:

$$\text{watermark}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} L_M(t, \omega) * \text{STFT}_p(t, \omega), \quad (14)$$

u kojoj je sa $\text{STFT}_p(t, \omega)$ označena STFT pseudo-slučajne sekvence p . Kreiran na osnovu funkcije L_M , watermark će biti prisutan samo u oblasti definisanosti te funkcije.

Nakon kreiranja adekvatnog watermark signala procedura njegovog umetanja u vremenskom domenu opisuje se relacijom:

$$s_w(t) = s(t) + \sigma * \text{watermark}(t), \quad (15)$$

gdje su sa $s(t)$ i $s_w(t)$ označeni originalni i watermarkirani govorni signal, respektivno. Parametar σ služi za podešavanje jačine umetnutog watermark-a.

3.1. Detekcija watermark-a

Detekcija umetnutog watermark-a omogućena je korišćenjem korelacione funkcije u vremenskom domenu. Naime, ukoliko se formira bilo koji novi signal, na isti način na koji je formiran watermark, za uspješnu detekciju neophodno je da bude zadovoljena relacija:

$$\sum_{t=t}^t s_w(t) * \text{watermark}(t) > \sum_{t=t}^t s_w(t) * f_i(t), \quad (16)$$

gdje, $f_i(t)$ predstavlja bilo koji signal kreiran kao i watermark, a $i=1, 2, 3, \dots$. Svrha kreiranja ovakvih signala je narušavanje autorskih prava, odnosno pokušaj dokazivanja vlasništva bez posjedovanja pravog ključa.

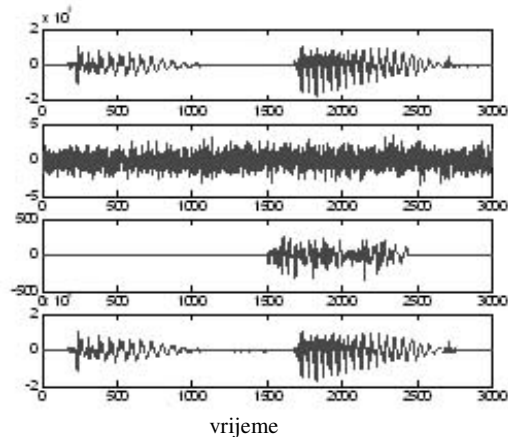
4. ILUSTRACIJA PROCEDURE KROZ PRIMJER

Pri praktičnoj implementaciji opisane procedure, korišćen je govorni signal maksimalne frekvencije $f_{max}=4\text{kHz}$. Polazna sekvenca p koja se u daljoj obradi filtrira vremenski-promjenljivim filtrom, definisana je kao bijeli Gausov šum.

STFT signala je računata koristeći prozor dužine 256 odbiraka, dok je za prikazivanje iste korišćen prozor sa 1024 odbirka. Za vrijednost praga ξ pri definisanju funkcije L_M korišćena je vrijednost:

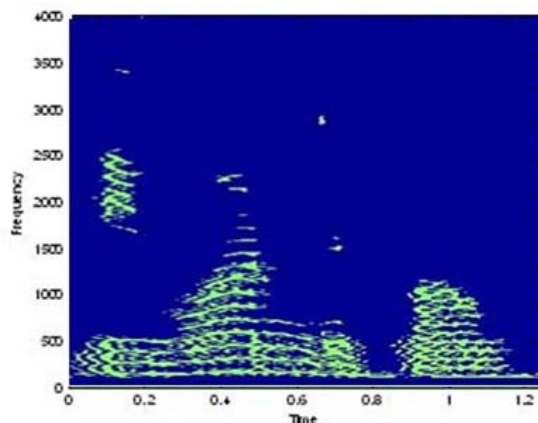
$$\xi = 5 * (\overline{\text{SPEC}}_s(t, \omega)).$$

Na Slici.2. prikazani su, redom, originalni govorni signal, Gausov šum, filtrirani Gausov šum, i signal sa watermark-om.



SI.2. Vremenski oblici govornog signala, Gausovog šuma, filtriranog Gausovog šuma-watermark-a, i watermarkiranog signala

Funkcija L_M korišćena pri vremenski-promjenljivom filtriranju ilustrovana je na Slici.3.



SI.3. Funkcija L_M

Detekcija je testirana korišćenjem 100 signala kreiranih na isti način kao i watermark. Slika.4. ilustruje da je u slučaju korišćenja watermark-a odziv detektora (središnji pik) višestruko veći nego u slučaju svih ostalih signala.

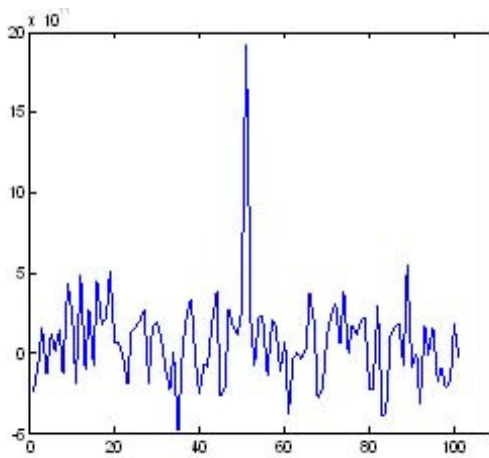


Fig.4. Odziv detektora

5. ZAKLJUČAK

U radu je prezentovana watermarking procedura zasnovana na vremenski-promjenljivoj filtriranju u vremensko-frekvencijskom domenu. Ovakav pristup omogućava efikasno modelovanje watermark-a u skladu sa regionom govornog signala u koji se umeće. Navedeni rezultati pokazuju efikasnost procedure, kako u perceptualnom pogledu tako i sa stanovišta detekcije.

REFERENCE

- [1] I.J.Cox, J.Kolian, F.T.Leigton, T.Shamoon, "Secure Spectrum Watermarking for Multimedia"(1996), *IEEE Trans.Image Processing*, Vol.6,pp.1673-1687.
- [2] H.J.Kim, "Audio Watermarking Techniques", *Pacific Rim Workshop on Digital Steganography*, Kyushu Institute of Technology, Kitakyushu, Japan, July, 2003.
- [3] M.F.Mansour, A.H.Tewfik, "Audio Watermarking by Time-Scale Modification", *International Conference on Acoustic, Speech, and Signal Processing*, 2001, Vol. 3, pp. 1353-1356.
- [4] M.F.Mansour, A.H.Tewfik, "Time-Scale Invariant Audio Data Embedding", *International Conference on Multimedia and Expo*, 2001.
- [5] LJ.Stanković, "Introduction to time-frequency analysis", *Time-Frequency Signal Analysis*, research monograph, 1993-2003.
- [6] S.Stanković, "About Time-Variant Filtering of Speech Signals with Time-Frequency Distributions for Hands-Free Telephone Systems", *Signal processing*, Vol.80, No.9, 2000.
- [7] G.Matz, F.Hlawatsch, W.Kozek, "Generalized Evolutionary Spectral Analysis and the Weyl Spectrum of Nonstationary Random Processes", *IEEE Transaction on Signal Processing*, Vol.45, No.6, Jun 1997, pp.1520-1534.