



Wytwarzanie znaku wodnego w czasie rzeczywistym

BORYS BOGDAN, JERZY ŁOPATKA

Wojskowa Akademia Techniczna, Wydział Elektroniki, Instytut Telekomunikacji,
00-908 Warszawa, ul. S. Kaliskiego 2

Streszczenie. Artykuł przedstawia jedną z metod generacji oraz opis funkcjonalny generatora znaku wodnego, działającego w czasie rzeczywistym. Generator ten jest jednym z elementów systemu znakowania sygnału akustycznego dla potrzeb transmisji danych w kanałach akustycznych. W artykule przedstawiono budowę generatora, sposób wytwarzania znaku wodnego oraz dodawania go do sygnału akustycznego z wykorzystaniem techniki DSP.

Słowa kluczowe: telekomunikacja, znak wodny, DSP, MPEG

Symbol UKD: 621.39

1. Wstęp

Znakowanie sygnału akustycznego (ang. *WaterMarking*) jest procesem polegającym na dodawaniu do tego sygnału pewnej dodatkowej informacji, którą chcemy zawrzeć w sygnale akustycznym. Oczywiście jest, że takie znakowanie ma sens wyłącznie wtedy, gdy jest percepcyjnie transparentne, czyli kiedy odbiorca (słuchacz) sygnału akustycznego nie jest w stanie usłyszeć zniekształceń sygnału akustycznego, spowodowanych znakowaniem.

Z drugiej strony, w zależności od przewidywanego zastosowania systemu znakowania sygnału akustycznego, należy wziąć pod uwagę problem odporności informacji zawartej w znaku wodnym na celowe lub przewidywalne manipulacje, np. dodawanie szumu, kompresję, filtrację, resampling, rekwantyzację, ucinanie, skalowanie itp. Ponadto, w zależności od przewidywanego zastosowania, należy dobrać odpowiednią przepływność binarną utworzonego kanału transmisji danych.

Ze względu na wykorzystanie znakowania sygnałów akustycznych możemy wymienić dwie zasadnicze kategorie znaków wodnych [1, 2]:

- tajne znaki wodne wykorzystywane do autoryzacji oraz w mechanizmach zapewniających integralność zawartości danych. Znak wodny jest tutaj poufnym kodem odczytywalnym tylko przez autoryzowane osoby mające dostęp do klucza, na podstawie którego informacja zawarta w znaku wodnym jest następnie deszyfrowana,
- publiczne znaki wodne funkcjonują jako nośniki informacji odczytywalnej przez dowolną osobę i pozwalają na przesyłanie informacji związanych z danym sygnałem akustycznym, a także na jego identyfikację.

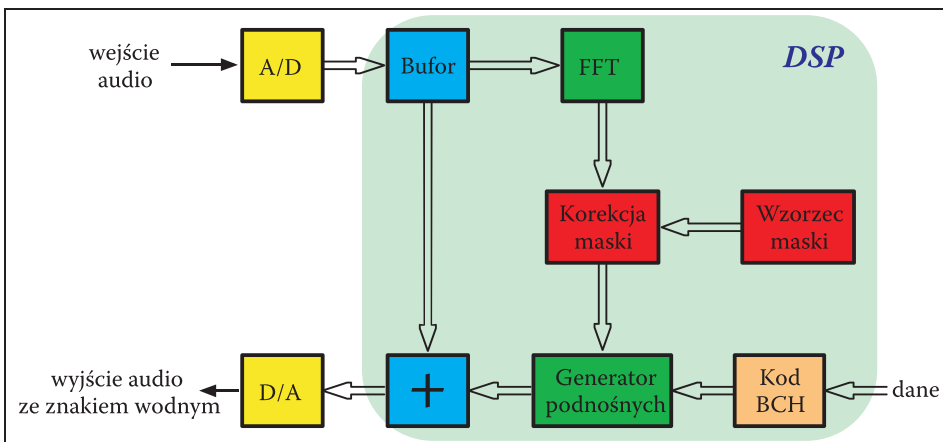
Obecnie stosuje się następujące podstawowe techniki znakowania wodnego sygnału audiofonicznego:

- kodowanie najmniej znaczących bitów (LBC) — dla sygnałów akustycznych transmitowanych w postaci cyfrowej,
- kodowanie fazowe,
- kodowanie z rozpraszaniem widma,
- maskowanie echa,
- maskowanie psychoakustyczne.

Opracowany generator wykorzystuje transmisję szerokopasmową z wykorzystaniem maskowania psychoakustycznego.

2. Zasada działania generatora znaku wodnego

Generator znaku wodnego został zrealizowany w sposób cyfrowy, z wykorzystaniem procesora sygnałowego DSP. Schemat funkcjonalny generatora został przedstawiony na rysunku 1. Elementami odpowiedzialnymi za przetwarzanie sygnału akustycznego z postaci analogowej na postać cyfrową i odwrotnie są ste-

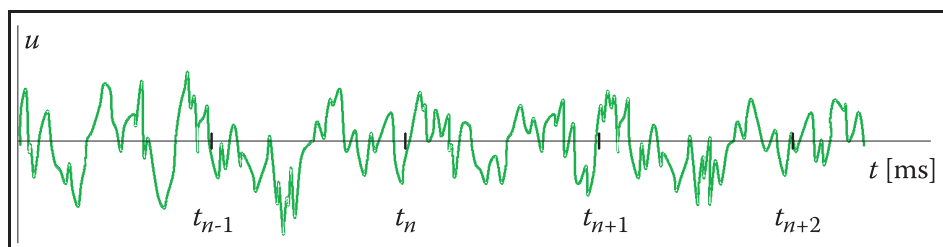


Rys. 1. Schemat funkcjonalny generatora znaku wodnego

reofoniczne, 24-bitowe przetworniki A/D oraz D/A, pracujące z częstotliwością próbkowania równą 48 kS/s. Uzyskany w wyniku przetwarzania A/D strumień danych rozdzielany jest następnie na dwa równoległe strumienie zawierające odpowiednio próbki sygnału z kanału lewego oraz prawego [3].

Oba uzyskane w powyższy sposób strumienie danych podawane są następnie do wspólnego bufora, pozwalającego na zgromadzenie niezbędnej liczby próbek dla potrzeb ich późniejszej analizy widmowej oraz w celu zsumowania ich z wygenerowanym znakiem wodnym (w dziedzinie czasu). Po zgromadzeniu zadanej liczby próbek realizowany jest proces obliczania tzw. maski znaku wodnego, tzn. dopuszczalnych poziomów składowych tego znaku w dziedzinie częstotliwości, dla których subiektywne odczucia przeciętnego słuchacza nie wykażą różnicy pomiędzy sygnałem wejściowym a sygnałem ze znakiem wodnym.

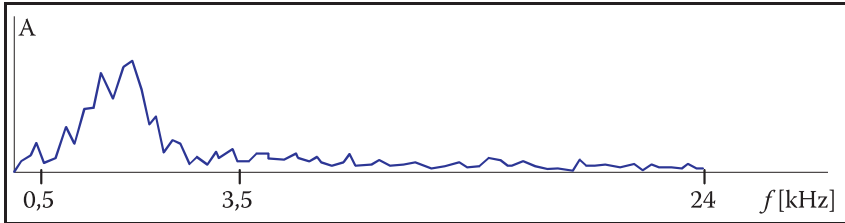
Wyznaczanie maski znaku wodnego realizowane jest w dziedzinie częstotliwości, zatem pierwszą operacją realizowaną na sygnale akustycznym jest wyznaczenie jego widma, ograniczonego zakresem częstotliwości wynikającym z częstotliwości próbkowania. Na rysunku 2 przedstawiono przykładowy przebieg analizowanego sygnału akustycznego w dziedzinie czasu, natomiast rysunek 3 przedstawia przykładowe widmo amplitudowe fragmentu sygnału z rysunku 2. Widmo chwilowe każdego z kanałów (lewego oraz prawego) wyznaczone jest w oparciu o algorytm FFT wykorzystujący arytmetykę zmiennoprzecinkową. Ponadto w celu zapewnienia „ciągłości” widmowej przetwarzanego sygnału wykorzystuje się tzw. nakładkowanie. Oznacza to, że wyznaczanie widma chwilowego odbywa się w oparciu o zestaw próbek zebranych w przedziale czasu $(t_{n+1}-t_n)$ oraz fragment zestawu próbek wykorzystanych w poprzednim przedziale czasu (t_n-t_{n-1}) , gdzie n oznacza kolejny moment analizy sygnału określony częstotliwością próbkowania oraz długością bufora analizy widmowej.



Rys. 2. Przykładowy przebieg sygnału akustycznego

Zasadniczym etapem wyznaczania maski znaku wodnego jest korekcja wzorca maski na podstawie wyznaczonych powyżej składników widma analizowanego sygnału akustycznego [4]. Z uwagi na fakt, że znacząca część energii sygnału akustycznego zawiera się w przedziale od 300 do 4000 Hz, ze względu na możliwe zniekształcenia amplitudowe na krańcach pasma rzeczywistych kanałów telekomu-

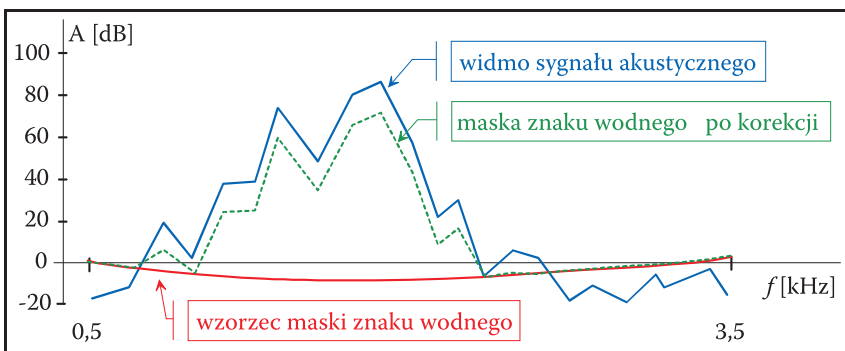
nikacyjnych, przeznaczonych do transmisji głosu, pasmo przydzielone dla znaku wodnego zostało ograniczone do zakresu od 500 do 3500 Hz. Z tego też względu do korekcji wzorca maski znaku wodnego wykorzystuje się wyłącznie fragment widma amplitudowego z rysunku 3, należący do powyższego zakresu częstotliwości.



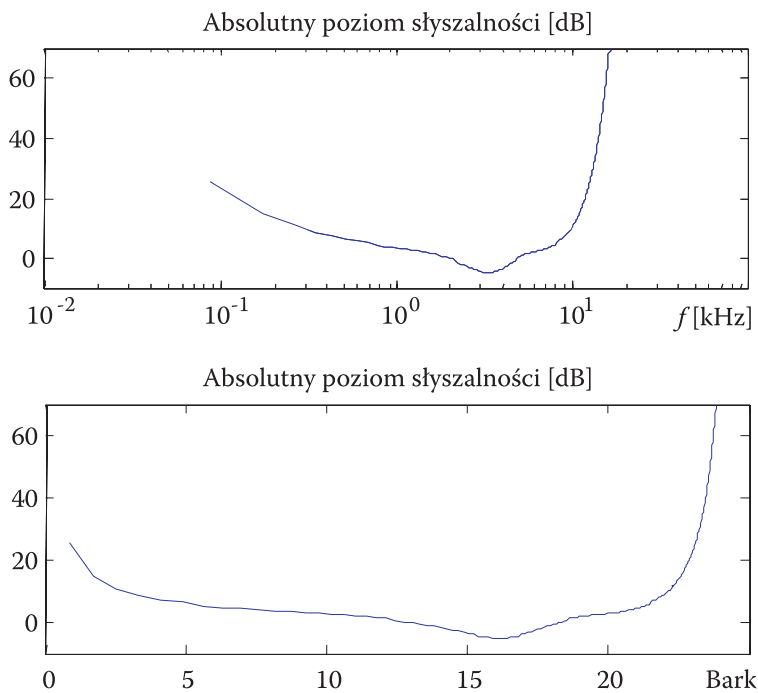
Rys. 3. Przykładowe widmo amplitudowe sygnału akustycznego

Tworzenie i korekcja maski znaku wodnego, którego ogólną zasadę przedstawiono na rysunku 4 wykorzystuje psychoakustyczny model czułości ludzkiego słuchu. Wykorzystany w niniejszym projekcie model psychoakustyczny bazuje na standardzie MPEG Layer 1. Operacja niezbędna do utworzenia maski znaku wodnego, adekwatnej do chwilowego widma amplitudowego sygnału akustycznego, odbywa się z częstotliwością 125 Hz, wynikającą z założenia, że częstotliwość próbkowania sygnału wynosi 48 kS/s oraz że wykorzystujemy blok próbek o długości 512 próbek (z nakładkowaniem 25%). Procedura, pozwalająca na utworzenie maski znaku wodnego, jest realizowana w kilku etapach.

Na etapie pierwszym wyznaczany jest wzorec maski znaku wodnego — tzw. absolutny poziom słyszalności. Wzorec ten (przedstawiony na rysunku 5) stanowi zbiór dopuszczalnych poziomów poszczególnych składowych znaku wodnego, dla których sygnał znaku wodnego nie jest słyszalny przez przeciętnego słuchacza. Wzorec ten został utworzony na podstawie psychoakustycznej analizy czułości ucha



Rys. 4. Korekcja maski znaku wodnego



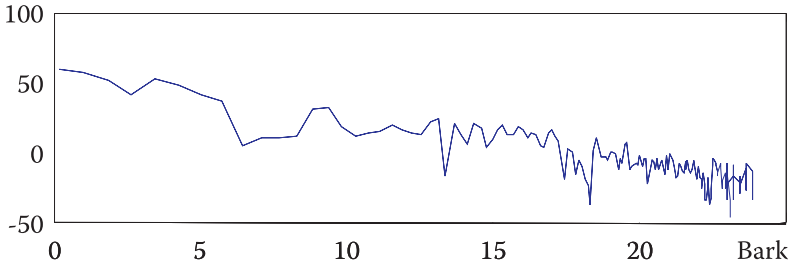
Rys. 5. Wzorzec maski znaku wodnego

ludzkiego z uwzględnieniem subiektywnych odczuć szerokiego grona słuchaczy. Stanowi on zatem zbiór (w dziedzinie częstotliwości) podstawowych poziomów dla poszczególnych składowych znaku wodnego, niezależny od chwilowego widma sygnału akustycznego.

W tym miejscu należy zauważyć, że wszystkie dalsze obliczenia związane z wyznaczaniem maski znaku wodnego będą realizowane w dziedzinie częstotliwości przekształconej do tzw. pasm Barka. Pasma Barka są to pasma (w dziedzinie częstotliwości), których szerokość wynika ze zdolności percepcyjnych ludzkiego słuchu, określonych na podstawie przyjętego modelu psychoakustycznego. Istotą tych pasm jest to, że obecność więcej niż jednej składowej o poziomie nie przekraczającym poziomu składowej maksymalnej występującej w tym paśmie nie jest dla statystycznego słuchacza odczuwalna. Przykładowo, w przyjętym modelu psychoakustycznym piąte pasmo zawiera się pomiędzy 300 a 400 Hz (szerokość 100 Hz), natomiast piętnaste pasmo zawiera się pomiędzy 2000 a 2320 Hz (szerokość 320 Hz).

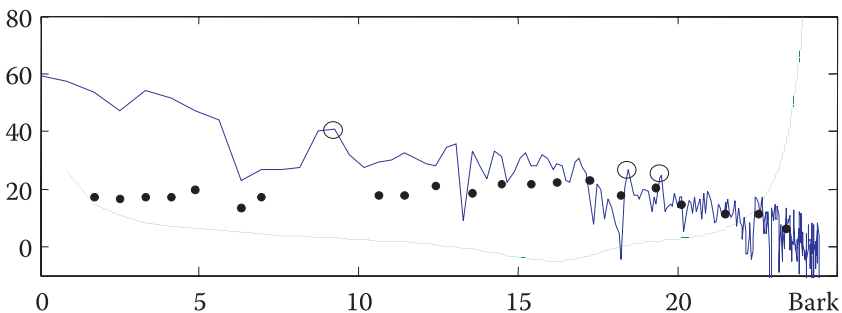
W drugim etapie wyznaczania maski znaku wodnego analizowany sygnał audio (w postaci cyfrowej) dzielony jest na segmenty (w dziedzinie czasu) po 384 próbki, co odpowiada 8 ms przy $f_p = 48$ kS/s. Odcinki te następnie przekształcane są do dziedziny częstotliwości za pomocą szybkiej transformaty Fouriera (FFT — *Fast Fourier Transform*). W celu ograniczenia przecieków widma oraz wpływu chwilowych ekstremów

każdy segment sygnału audio dopełniany jest do długości 512 próbek za pomocą 128 próbek z poprzedniego segmentu (nakładkowanie), a następnie okienkowany jest z wykorzystaniem okna Hanninga. W wyniku tego uzyskuje się chwilowe widmo amplitudowe sygnału audio, co przedstawiono na rysunku 6.



Rys. 6. Przykładowe widmo sygnału audio

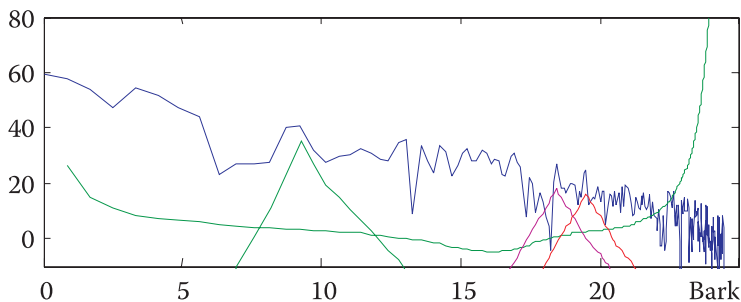
W ramach kolejnego etapu analizowany sygnał (w dziedzinie częstotliwości) dzielony jest na pasma (tzw. pasma Barka), wynikające z zastosowanego modelu psychoakustycznego. Następnie każde pasmo Barka klasyfikowane jest jako „tonowe” lub „szumowe”. Oznacza to, że dla każdego pasma określane jest, czy w danym paśmie decydującą składową jest składowa tonalna czy też szumowa. Kryterium przyjęte do oceny danego pasma sprowadza się do wyznaczenia różnic pomiędzy maksymalnym poziomem składowej sygnału audio (FFT) w środku danego pasma Barka a poziomami jego czterech sąsiednich składowych (dwóch powyżej i dwóch poniżej). Jeżeli wyznaczone różnice przekraczają odpowiednie wartości progowe, to wówczas dane pasmo klasyfikowane jest jako „tonowe”, natomiast w przeciwnym wypadku — jako „szumowe”. Przykład oznaczania poszczególnych pasm Barka jako „tonowe” oraz „szumowe” pokazano na rysunku 7. Linia przerywaną zaznaczono wzorzec maski znaku wodnego (patrz rysunek 5).



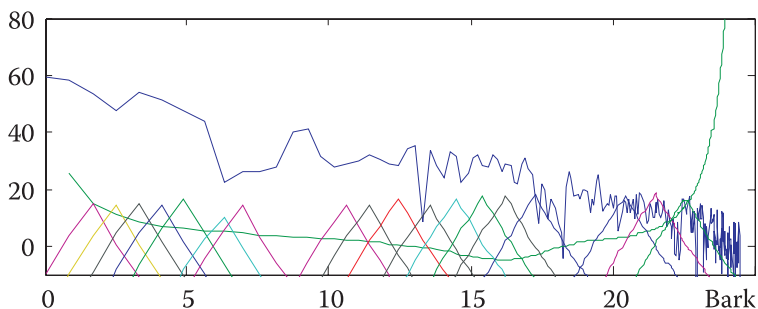
Rys. 7. Przykład klasyfikacji widma sygnału audio na składowe „tonowe” (kółka) oraz „szumowe” (kropki)

Równocześnie z klasyfikacją pasm odbywa się wyznaczenie maksymalnego poziomu sygnału dla każdego z tych pasm, przy czym z uwagi na różną czułość ludzkiego słuchu dla składowych „tonowych” oraz „szumowych” (występujących w danym paśmie) poziom ten jest odpowiednio wyższy w przypadku składowych „tonowych” oraz niższy dla składowych „szumowych”, co pokazano na rysunku 7.

W ramach piątego etapu wyznaczane są krzywe maksymalnych poziomów znaku wodnego dla każdego z pasm Barka oraz pasm sąsiednich (tzw. funkcje rozpraszające), z uwzględnieniem podstawowej składowej występującej w danym paśmie. Z uwagi na różną czułość ludzkiego słuchu w danym paśmie, w zależności od dominującej składowej („tonowa” lub „szumowa”) funkcje rozpraszające dla składowej „tonowej” charakteryzują się odpowiednio wyższymi poziomami niż funkcje rozpraszające dla składowej „szumowej”. Przykłady funkcji rozpraszających dla składowych „tonowych” oraz „szumowych” przedstawiają odpowiednio rysunki 8 i 9.



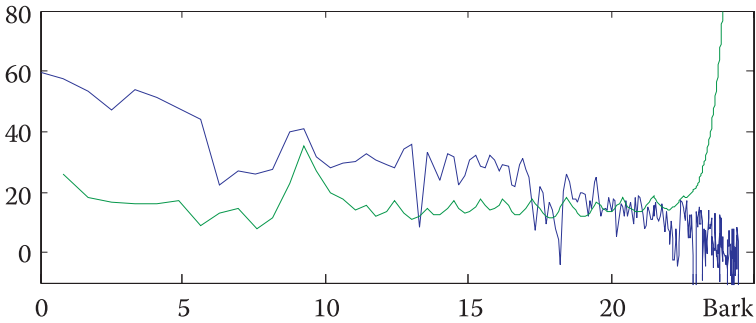
Rys. 8. Funkcje rozpraszające dla składowych „tonowych”



Rys. 9. Funkcje rozpraszające dla składowych „szumowych”

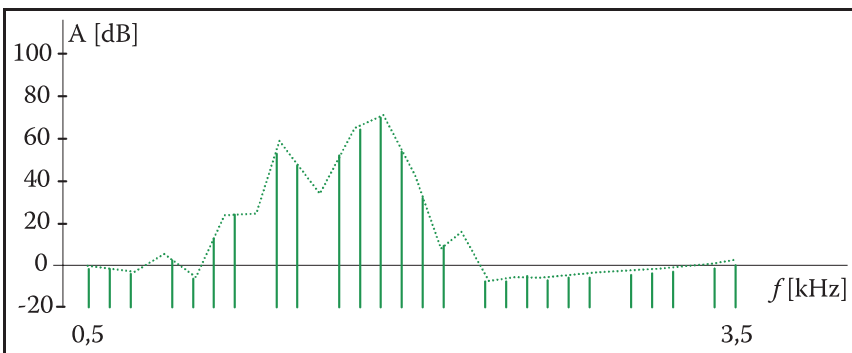
W oparciu o wzorec maski (absolutny poziom słyszalności) oraz wyznaczone uprzednio krzywe maksymalnych poziomów znaku wodnego w poszczególnych podpasmach (funkcje rozpraszające) wyznaczana jest globalna maska znaku wodnego. Maska ta obliczana jest poprzez zsumowanie maksymalnych poziomów wszystkich ww. trzech składowych w dziedzinie częstotliwości. Stanowi ona zarazem

ograniczenie (w dziedzinie częstotliwości) poziomów poszczególnych składowych znaku wodnego, które zapewni, że po dodaniu do sygnału audio sygnału znaku wodnego u przeciętnego słuchacza nie zostanie zaobserwowane pogorszenie jakości sygnału pierwotnego.



Rys. 10. Globalna maska znaku wodnego

Właściwy znak wodny, przedstawiony w dziedzinie częstotliwości na rysunku 11, generowany jest na podstawie wyznaczonej uprzednio maski. W zrealizowanym układzie odbywa się to poprzez wytworzenie odpowiednich składowych harmonicznych metodą OFDM (ang. *Orthogonal Frequency Division Multiply*), w oparciu o wyznaczoną uprzednio maskę oraz z uwzględnieniem transmitowanej informacji. Z oczywistych względów przesyłana za pomocą znaku wodnego informacja jest dodatkowo zabezpieczana przed błędami poprzez zastosowanie kodowania korekcyjno-detekcyjnego kodem spletowym BCH. Na ostatnim etapie, wygenerowany w ten sposób znak wodny sumowany jest z sygnałem akustycznym, odpowiednio w kanale lewym oraz prawym, a następnie przetwarzany do postaci pierwotnej za pomocą przetwornika D/A.



Rys. 11. Widmo przykładowego sygnału znaku wodnego

3. Wnioski

Przedstawiony powyżej generator znaku wodnego został wykonany dla potrzeb transmisji danych w kanale, w którym pierwotnie przesyłana jest informacja w postaci sygnału akustycznego. Rozwiązanie takie umożliwia między innymi wymianę danych w sposób niezauważalny dla słuchacza, a więc bez pogarszania subiektywnych odczuć dotychczasowych użytkowników komunikujących się za pomocą sygnałów akustycznych. Istotnym mankamentem takiego rozwiązania jest jednak niewielka przepływność binarna, ograniczająca potencjalne obszary zastosowania.

Istotnymi problemami wymagającymi dalszych badań są: dostosowanie algorytmów odtwarzania synchronizacji po stronie odbiorczej, zaproponowanych w [5] do kanałów cyfrowych oraz kompensacja zniekształceń wprowadzanych przez cyfrowe kanały z kompresją danych. Prace w tym kierunku są kontynuowane.

Artykuł wpłynął do redakcji 19.07.2006 r. Zweryfikowaną wersję po recenzji otrzymano 24.10.2006 r.

LITERATURA

- [1] P. GAJEWSKI, J. ŁOPATKA, Z. PIOTROWSKI, *New model of correspondent verification in telecommunication digital channels*, Materiały konferencyjne CD-ROM, Signal Processing Symposium, Wilga, 3-5.06.2005.
- [2] P. GAJEWSKI, J. ŁOPATKA, Z. PIOTROWSKI, *New model of identity checking in telecommunication digital channels*, 7th NATO Regional Conference on Military Communications and Information Systems 2005, 04-05.10.2005, Conference Proceedings, Zegrze, 126-131.
- [3] W. BUTRYN, *Dźwięk cyfrowy*, WKŁ, Warszawa, 2002.
- [4] T. PAINTER, A. SPANIAS, *Perceptual Coding of Digital Audio*, Proceedings of the IEEE, vol. 88, no. 4, April 2000.
- [5] P. GAJEWSKI, J. ŁOPATKA, Z. PIOTROWSKI, *A new method of frequency offset correction using coherent averaging*, Journal of Telecommunications and Information Technology, no. 1, 2005, 142-146.

B. BOGDAN, J. ŁOPATKA

Watermark generation in real time

Abstract. The paper presents a method of watermark signal generation and functional description of a real time watermark signal generator. This generator was designed for transmission of a hidden signature over transmission channels. Generator structure, inner structure of a watermark signal, and the used model of psycho-acoustic mask are described.

Keywords: telecommunication, watermarking, DSP, MPEG

Universal Decimal Classification: 621.39

