

Analiza widmowa sygnałów mowy przy pomocy komputera PC

Wzrost wydajności komputerów PC umożliwił stosowanie ich w aplikacjach zarezerwowanych dotychczas dla systemów specjalizowanych. Jednym z takich obszarów jest widmowa analiza sygnałów.

W artykule przedstawiamy interesujące rozwiązanie tego problemu przy pomocy typowego komputera PC, karty przetwornika A/C i specjalnie przygotowanego oprogramowania.

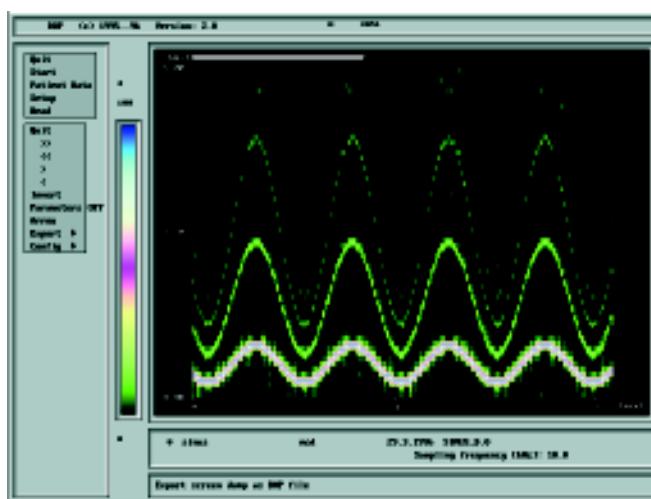
Mianem sygnału najczęściej określamy zamienioną na sygnał elektryczny informację towarzyszącą określonemu zjawisku. Może to być np. przebieg ciśnienia, drgań, prędkości, dane z zakresu ekonomii lub biologii, a także, oczywiście, przebiegi elektryczne. Sygnały na ogół mają charakter ciągły i najczęściej są funkcjami czasu. Sygnały występują w roli nośników informacji w wielu różnych dziedzinach - np. technice, biologii, ekonomii, fizyce, astronomii i medycynie. Istnieje wiele możliwości klasyfikowania sygnałów - ze względu na rodzaj energii, rodzaj informacji, losowy bądź deterministyczny charakter. Zainteresowanych szerszym potraktowaniem problematyki klasyfikacji sygnałów odsyłamy do literatury [1]. Informacja niesiona przez sygnał może być przedstawiona w dwojaki sposób - w dziedzinie

formaty Fouriera (FFT) umożliwiło dynamiczny rozwój analizy widmowej i jej zastosowań [2]. Obecnie podstawowym sposobem określania rozkładu częstotliwościowego jest metoda bezpośredniej transformacji Fouriera, wykorzystująca algorytm FFT, realizowana przez specjalizowane układy, procesory sygnałowe (DSP) lub komputery ogólnego przeznaczenia. Pojawienie się w latach 80-tych procesorów DSP stanowiło kolejny czynnik stymulujący rozwój zastosowań analizy widma. Procesory te miały nad wczesnymi komputerami klasy PC przewagę w postaci wyższej częstotliwości zegara oraz specjalizowanego do zastosowań w przetwarzaniu sygnałów zestawu instrukcji, które np. umożliwiały wykonanie mnożenia i dodawania w jednym cyklu (mnożenie i dodawanie to podstawowe operacje w przetwarzaniu

sygnałów), czy adresowanie kołowe dla potrzeb FFT. Od wielu lat produkowane są karty rozszerzeniowe do komputerów PC, zawierające procesory sygnałowe z niezbędnymi dodatkowymi układami. Okazuje się jednak, że ostatnio producenci kart DSP do komputerów PC przestają rozwijać te produkty. Dlaczego - przyczyną jest lawinowo wzrastająca szybkość procesorów używanych w komputerach PC - 8 lat temu PC AT o częstotliwości zegara 8MHz stanowiło szczyt możliwości - dzisiaj procesory Pentium - z wbudowanym dobrym procesorem zmiennoprzecinkowym - osiągają 200MHz. Jeśli więc myślimy o sprzęcie stacjonarnym, przemawiający

dotąd za DSP argument w postaci wyższej częstotliwości zegara znika. Miejsce DSP zaczyna zajmować komputer PC, który w pewnym zakresie wartości częstotliwości próbkowania i przy umiarkowanych wymaganiach, jeśli chodzi o sposób prezentacji rezultatów analizy w czasie rzeczywistym, daje zupełnie zadowalające wyniki.

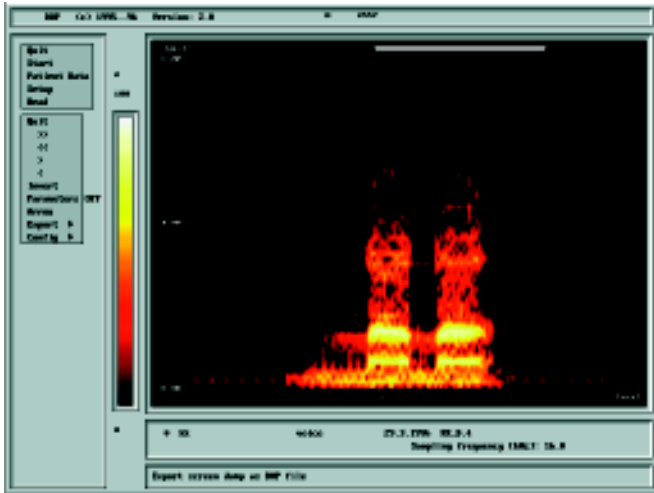
W wielu praktycznych sytuacjach utrudnienie stanowi fakt, że własności źródła sygnałów zmieniają się w czasie, a więc własności powstających sygnałów także ulegają zmianom w funkcji czasu, w niektórych przypadkach w losowy sposób. Narzuca to konieczność przedstawienia sygnału na płaszczyźnie, w funkcji dwóch zmiennych - czasu



Rys.1.

czasu i w dziedzinie częstotliwości. Pierwsza z w/w możliwości to po prostu przebieg sygnału w funkcji czasu, w przypadku złożonego kształtu bardzo trudny w interpretacji. Druga z tych możliwości to prezentacja rozkładu częstotliwościowego (widmowego, widmowej gęstości mocy). Rozkłady częstotliwościowe sygnałów wyznaczano początkowo m.in. przy pomocy zestawów filtrów lub analizatorów tzw. heterodynowych, później natomiast zaczęto stosować metody komputerowe. Czynnikiem hamującym rozwój przetwarzania sygnałów, w tym także analizy widmowej, była niedostateczna efektywność algorytmów obliczeniowych, i dopiero opracowanie w roku 1965 algorytmu szybkiej trans-

i częstotliwości tak, aby ewolucja struktury widmowej w czasie była dobrze widoczna. Na potrzeby takich właśnie prezentacji powstało wiele metod np. spek-



Rys.2.

trogram, prezentacja Wigner-Ville, transformacja falowa. Najszerzej stosowany jest najwcześniej powstały spektrogram, będący wynikiem zastosowania transformacji Fouriera i podniesienia do kwadratu jej produktów w odniesieniu do krótkich, następujących po sobie sekwencyjnie odcinków sygnału. Pojęcie „krótki” nie jest dobrze określone i oznacza, że odcinki sygnału powinny być na tyle krótkie, by wspomniana ewolucja struktury widmowej była widoczna w kolejnych rozkładach widmowej gęstości mocy, a jednocześnie nie powinno to powodować konsekwencji w postaci zniekształceń widma, znanych pod nazwą rozmycia lub przecieku. Spektrogram może być prezentowany w różny sposób - pseudotrójwymiarowo lub na płaszczyźnie z wykorzystaniem skali szarości lub barw.

Analizator widma działający według algorytmu FFT zaimplementowany został na komputerze PC Pentium 100MHz. W komputerze zainstalowano 16-kanalową kartę we/wy analogowych o rozdzielczości 12 bitów, działającą z częstotliwościami próbkowania do 100kHz (w pojedynczym kanale), oraz kartę graficzną Paradise Bali 1MB. Analizator umożliwia m.in. dobór parametrów akwizycji i analizy, jak częstotliwość próbkowania, długość ciągu poddawanego transformacji FFT, długość transformaty (istnieje możliwość uzupełnienia ciągu próbek pewną liczbą zer). Uzyskiwane szybkości pracy jednokanałowej (częstotliwości próbkowania) z jednoczesną prezentacją sięgają przy stosunkowo krótkich oknach danych (rzędu 10ms) około 30kHz. Wydłużenie okna danych umożliwi pracę z wyższymi częstotliwościami próbkowania. Należy zwrócić uwagę, że analizator prezentuje wyniki w czasie rzeczywistym w trybie wysokiej rozdzielczości (1024*768 pikseli), co stanowi także znaczne obciążenie obliczeniowe. Rezygnacja ze współbieżnej

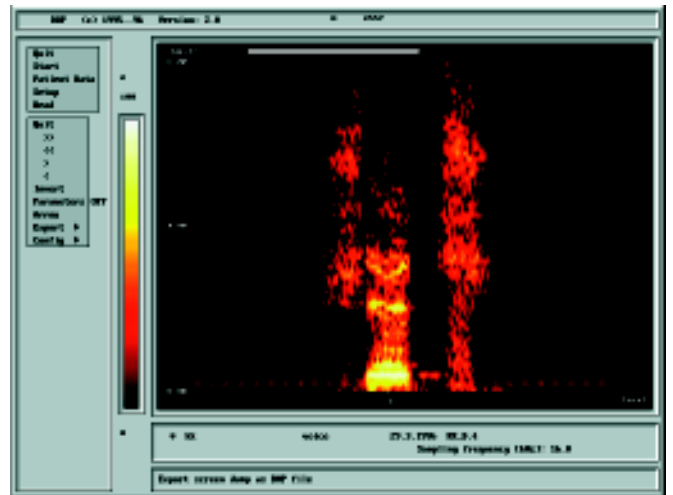
z obliczaniem widma prezentacji daje około 60% wzrost szybkości działania analizatora. Oczywiście ograniczenie rozdzielczości do standardu VGA zapewniłoby znaczne przyspieszenie wyrowadzenia wyników analizy widmowej na ekran, a więc umożliwiłoby stosowanie jeszcze wyższych częstotliwości próbkowania.

Rys.1 przedstawia spektrogram próbkowanego z częstotliwością 10kHz sygnału sinusoidalnego o częstotliwości nieco powyżej 700Hz, z modulacją częstotliwości sygnałem o częstotliwości 2Hz. Długość okna danych wynosiła tu 128 próbek. Na ekranie wyraźnie widoczne są składowa podstawowa, 3, 5 i 7 harmoniczna. Zaobserwować można także skutki tzw. nakładania się widm (aliasing), czyli efekty niespełnienia twierdzenia o próbkowaniu dla harmonicznych sygnału począwszy od 7-tej wyżej.

Rys.2 i rys.3 przedstawiają natomiast wyniki analizy widmowej sygnału mowy. Konkretnie są to wyrazy „mama” i „sieć”. Sygnał poddano próbkowaniu z częstotliwością 16kHz, natomiast długość okna danych wynosi 128 próbek - czyli rozdzielczość czasowa analizy wynosi 8ms. Wyraźnie widoczne są w przypadku głosek dźwięcznych „a” oraz „e” częstotliwości rezonansowe toru głosowego, tzw. formanty. Formantów jest najczęściej 4 lub 5, a ich częstotliwości zależne są od budowy anatomicznej toru głosowego. W przypadku głosek bezdźwięcznych „ś” i „ć” występują szerokie widma o charakterze szumowym. Pierwsza głoska „m” wyrazu „mama” trwa znacznie dłużej niż druga, choć zapewne wymawiając ten wyraz nikt nie zdaje sobie z tego sprawy. Pionowe „smużki” wyraźnie widoczne w niektórych miejscach sonogramu wyrazu „mama” czy głoski „e” w wyrazie „sieć” są wynikiem obecności tzw. tonu krztaniowego, czyli drgań więzadeł (strun) głosowych, który stanowi pobudzenie toru głosowego. Częstotliwość tonu krztaniowego jest dosyć niska - waha się od kilkudziesięciu do kilkuset Hz,

a można ją określić na podstawie takich widm przy pomocy pewnych specjalnych metod przetwarzania sygnałów. Wyznaczenie własności częstotliwościowych sygnału mowy - formantów, tonu krztaniowego, ewolucji formantów w czasie - stanowi pierwszy krok w przetwarzaniu tego sygnału, jego transmisji czy rozpoznawaniu. Przetwarzanie sygnału mowy jest dziedziną od dłuższego czasu intensywnie rozwijającą się i obejmuje bardzo szeroki zakres problematyki - od anatomii przez przetwarzanie sygnałów po lingwistykę oraz systemy eksperckie. Zainteresowanych tą tematyką odsyłamy do literatury [3, 4].

Wracając zaś do samego analizatora - stanowi on narzędzie pozwalające na analizę widmową sygnałów akustycznych i ponadakustycznych - i to nie tylko pochodzących z generatorów czy będących wypowiedziami własnymi lub członków rodziny, ale także sygnałów występujących przy testowaniu sprzętu audio (np. ewolucja odpowiedzi głośnika na pobudzenie) i wielu, wielu innych - wszystko zależy od wyobraźni użytkownika! Programowanie, przy użyciu którego uzyskano zaprezentowane wyniki, może zostać zaimplementowane także na komputerze wyposażonym w procesor Am586 lub DX4



Rys.3.

100MHz, i będzie działał w czasie rzeczywistym z niższymi częstotliwościami próbkowania. Możliwości wykorzystania oprogramowania będą oczywiście rosły ze wzrostem wydajności komputerów PC i kart graficznych.

Krzysztof Kałużyński
Krzysztof Mikołajczyk

Bibliografia

1. Ozimek E. *Podstawy teoretyczne analizy widmowej sygnałów*, PWN, 1985
2. Oppenheim A.V., Schaffer R.W. *Cyfrowe przetwarzanie sygnałów*, WKiŁ 1979
3. Tadeusiewicz R. *Sygnał mowy*, WKiŁ, 1988
4. Jassem W. *Podstawy fonetyki akustycznej*, PWN, 1973