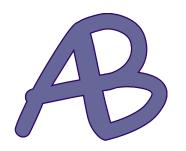
Architektura systemów komputerowych Ćwiczenie 3

Karta dźwiękowa

- Cyfrowy zapis dźwięku
 - Digitalizacja dźwięku
 - kompresja
- Budowa i działanie karty dźwiękowej
 - ewolucja kart dźwiękowych
 - Budowa karty dźwiękowej (schemat)
 - Synteza FM
 - Synteza Wavetable
 - Procesor DSP
- Dźwięk przestrzenny
 - Systemy dźwięku przestrzennego
 - Dolby Digital



Część 1



Cyfrowy zapis dźwięku

Trochę fizyki

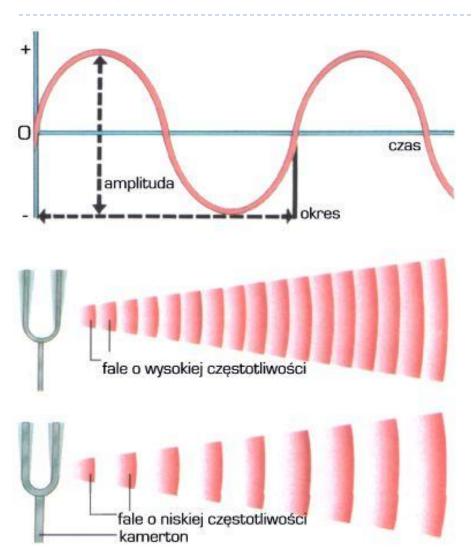


- ✓ Dźwięk jest falą rozchodzącą się w gazach i cieczach. Fala ta przenoszona jest poprzez cykliczne zmiany ciśnienia ośrodka.
- ✓ Ze względu na zakres częstotliwości można rozróżnić cztery rodzaje tych fal:
 - infradźwięki poniżej 20 Hz,
 - dźwięki słyszalne 20 Hz 20 kHz słyszy je większość ludzi,
 - ultradźwięki powyżej 20 kHz,
 - hiperdźwięki powyżej 10^10 Hz.

I. Cyfrowy dźwięk

Trochę fizyki





Drgania przełożyć można bezpośrednio na zmienny w czasie prąd elektryczny.

Wykres zmian prądu w czasie będzie odpowiadał wykresowi zmian ciśnienia ośrodka w punkcie, w którym znajduje się odbiornik.

B

Trochę fizyki

Dźwięk jako zjawisko fizyczne opisują trzy główne cechy:

- częstotliwość (podawana w Herzach (Hz) określa liczbę drgań na sekundę, np. 440 Hz oznacza 440 drgań na sekundę) — odbieramy ją jako wysokość dźwięku;
- natężenie dźwięku (podawane w watach na metr kwadratowy (W/m2) — w praktyce stosowana jest jednak inna miara: poziom natężenia dźwięku, którego jednostką jest decybel — dB);
- widmo dźwięku (zawartość składowych harmonicznych dźwięku, czyli to, co stanowi o charakterystycznym brzmieniu, barwie np. instrumentu).

Barwa dźwięku – subiektywna cecha dźwięku, która pozwala odróżnić brzmienia różnych instrumentów lub głosu. Uzależniona jest od ilości, rodzaju i natężenia tonów składowych (głównie od jego widma od pratur Bartoszewski - Informatyka i systemy informatyczne, sem. 1- WYKŁAL czestotliwości).

Kwantyzacja



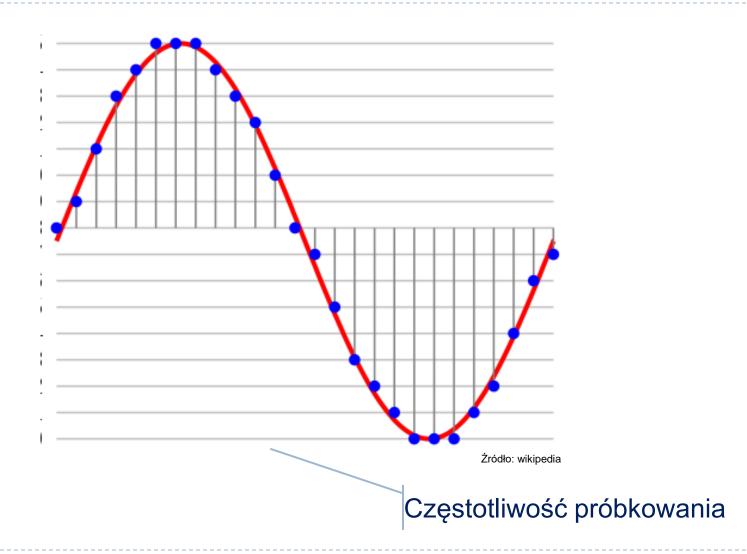
Dane dźwiękowe pozyskiwane są zwykle z postaci analogowej. Następnie przekształcane są postać cyfrową - digitalizacja

Digitalizacja dźwięku przebiega w trzech etapach:

- 1. próbkowanie,
- 2. kwantowanie,
- 3. kodowanie.

B

Cyfrowy zapis dźwięku - PCM





Cyfrowy zapis dźwięku - PCM

Częstotliwością próbkowania nazywamy ilość odczytów wartości sygnału wykonanych w ciągu sekundy. Wyprażmy ją w hercach (Hz).

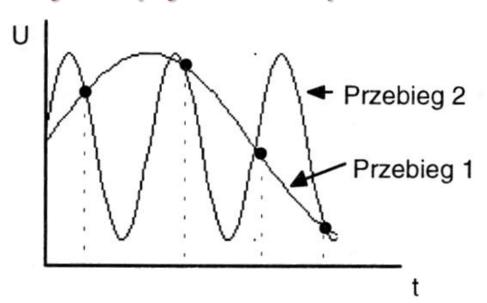
Rozdzielczością próbkowania (ilością poziomów kwantyzacji) nazywamy ilość stanów które przyjąć może próbkowany sygnał. Wyrażamy ją w bitach (ilość bitów potrzebnych do zapisania jednej próbki).

PCM ang. Pulse Code Modulation - modulacja kodowoimpulsowa - to najpopularniejsza metoda reprezentacji sygnału analogowego w systemach cyfrowych.

B

Cyfrowy zapis dźwięku - PCM

Jaka jest optymalna częstotliwość próbkowania?



Zjawisko to opisuje kryterium Nyquista:

✓ do poprawnego odwzorowania sygnału wystarczy poddać go próbkowaniu z dwa razy większą częstotliwością.

Źródło: Wojtuszkiewicz K., Urządzenia techniki komputerowej, Mikom



Dźwięk – pasmo przenoszenia

Dźwięk w zakresie słyszalnym przez człowieka:

✓ ok. 20 Hz – 20 kHz

Mowa:

- ✓ szerokopasmowy sygnał mowy: 50 Hz 7 kHz
- √ wąskopasmowy sygnał mowy: 200 Hz 3,2 kHz

Oznacza to, że jeśli najwyższą częstotliwością, jaką możemy usłyszeć, jest 20 kHz (teoretyczny zakres ludzkiego słuchu to 16 Hz - 20 kHz), to wystarczy, że sygnał zostanie próbkowany z częstotliwością 40 kHz. W praktyce jednak stosowany jest standard 44.1 kHz (w przy-padku płyt audio).



Dźwięk – pasmo przenoszenia

Częstotliwość próbkowania

Zastosowania

8,0 kHz	Telekomunikacja
32,0 kHz	Cyfrowe radio i telewizja
44,1 kHz	Dyski CD Audio
48,0 kHz	Magnetofony DAT, telewizja HDTV
96,0 kHz	Dyski DVD Audio



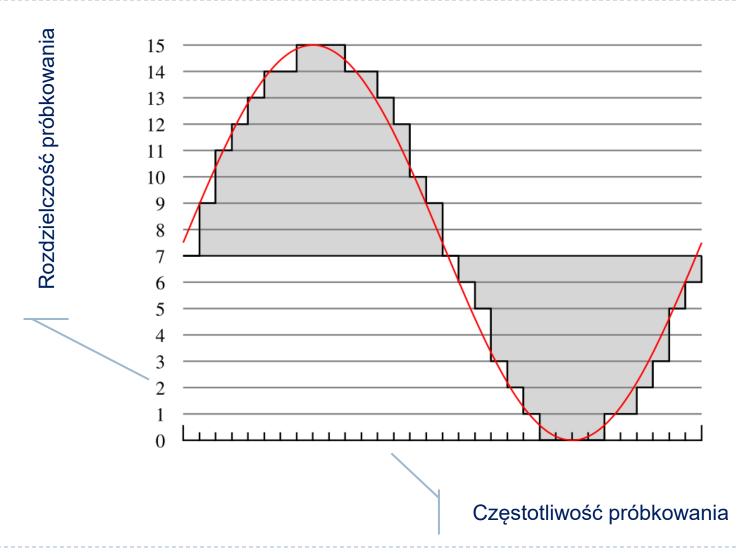
Digitalizacja dźwięku

Drugim etapem digitalizacji dźwięku jest kwantowanie

Kwantowanie polega na:

- podzieleniu ciągłego zbioru wartości sygnału skończoną liczbę sąsiadujących ze sobą przedziałów i ustaleniu poziomów kwantowania (tj. określonych wartości każdego przedziału, reprezentujących wszystkie wartości w tym przedziale),
- następnie przypisaniu każdej próbce odpowiedniego (najbliższego) poziomu kwantowania;

Cyfrowy zapis dźwięku - PCM



Cyfrowy zapis dźwięku - PCM

Kodowanie sygnału - ostatni etap procesu przetwarzania sygnału analogowego na postać cyfrową polega na przyporządkowaniu poszczególnym wartościom (poziomom kwantowania) odpowiednich wartości cyfrowych (tzw. słów kodowych).

W przypadku płyty CD mamy do czynienia z 16-bitowym słowem (65 tyś progów kwantyzacji). Spotykane obecnie urządzenia mają możliwość zapisywania sygnału także z 24- i 32-bitową rozdzielczością.

WAVE



WAV(WAVE) to najpopularniejszy format zapisu plików audio — bez utraty jakości (przy zachowaniu wysokiej częstotliwości próbkowania i rozdzielczości). Pliki te posiadają rozszerzenie wav.

Podstawową zaletą tych plików jest ich jakość. Jedyną wadą jest — duża objętość, zależna od częstotliwości próbkowania i rozdzielczości (np. jedna minuta nagrania przy 44.1 kHz/16-bit zajmie ok. 10 MB).

Pliki tego formatu możemy bezpośrednio nagrać na płytę CD i odtworzyć na domowym zestawie audio.

AIFF to drugi obok WAV format, który zapewnia wysoką jakość dźwięku. Używany zarówno na platformach PC, jak i Macintosh. Podobnie jak WAV nie wykorzystuje kompresji dźwięku. Pliki tego typu noszą rozszerzenie .aif.

Bezstratna kompresja dźwięku

FLAC (ang. Free Lossless Audio Codec) - format bezstratnej kompresji dźwięku z rodziny kodeków Ogg.

FLAC zapewnia kompresję 30 - 50%. W przypadku tego formatu pliki noszą rozszerzenie ".flac".



W przeciwieństwie do stratnych kodeków dźwięku takich jak MP3 i AAC, kodek FLAC nie usuwa żadnych danych ze strumienia audio, dzięki czemu po dekompresji otrzymujemy dźwięk identyczny z pierwowzorem.

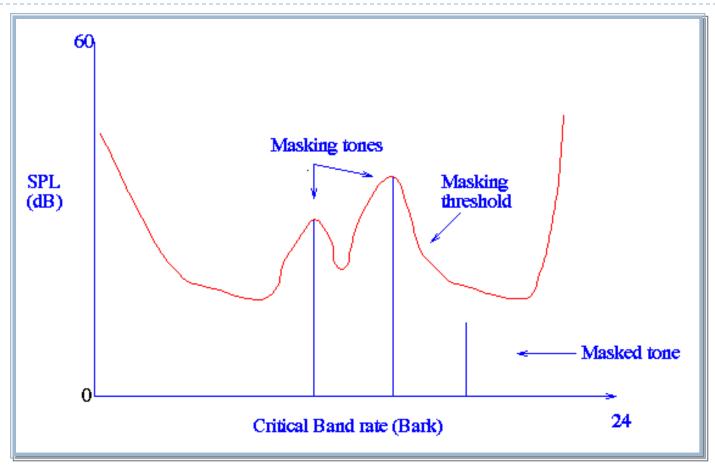
Stratna kompresja dźwięku

Kompresja stratna to metody zmniejszania ilości bitów potrzebnych do wyrażenia danej informacji, które nie dają gwarancji, że odtworzona informacja będzie identyczna z oryginałem.

Kompresja stratna jest możliwa ze względu na sposób działania ludzkich zmysłów.

- Algorytmy kompresji stratnej zazwyczaj posługują się modelami psychoakustycznymi, psychowizualnymi itd., aby odrzucić najmniej istotne dane o dźwięku, obrazie, pozostawiając dane o wyższej wartości dla rozpoznawania tej informacji (akustycznej, wizualnej) przez zmysły.
- Ilość odrzucanych danych jest zazwyczaj określana przez stopień kompresji.

Model psychoakustyczny



Zjawiskami słuchowymi najczęściej wykorzystywanymi przy opracowywaniu modeli psychoakustycznych są: maskowanie dźwięków, percepcja wysokości dźwięków



Kompresja MPEG (dla dźwięku)

Format MEPG Audio (Moving Picture Experts Group - Audio) wykorzystano kodowanie pychoakustyczne. Pozwala to na usunięcie dużej części informacji przy małej zmianie percepcji dźwięku.

Zasada kodowania dźwięku w formacie MPEG:

- 1. Za pomocą 32 filtrów pasmowych przeprowadza się analizę spektrum częstotliwości z podziałem na subpasma.
- 2. Sygnały o częstotliwościach poza zakresem słuchu człowieka są ignorowane.
- 3. Usuwane są sygnały maskowane.
- 4. Sygnał kodowany jest matematycznie (kodowanie Hufmana)
 - próbka może być kodowana na różnej długości słowa.

Kompresja MPEG (dla dźwięku)

Warstwa

Przepływność dla częstotliwości 44,1 kHz

I	256–384 kb/s
II	192–256 kb/s
III	128–192 kb/s



Kompresja MPEG (dla dźwięku)

Warstwę I utworzono do zastosowań na kasetach DCC (Digital Compact Cassete – cyfrowa kaseta kompaktowa). Oferuje ona stały stopień kompresji 1:4, czyli 384 kb/s przy 44,1 kHz.

Warstwa II operuje strumieniami wyjściowymi 192-286 kb/s, czyli stopniami kompresji 1:8-1:6. Znajduje ona zastosowanie w radiofonii cyfrowej i telewizji cyfrowej – obecnie wyparta przez standard MPEG 2 (taki jak na DVD)

Warstwa III - (MP3) - była początkowo opracowana z myślą o pracy w zakresie 112-128 kb/s jednak ze względu na dużą elastyczność algorytmu można uzyskać zakres do 64 kb/s do 384 kb/s.

I. Cyfrowy dźwięk

Porównanie różnych kodeków





Widma wybranych kodeków



Kompresja MPEG 2 (dla dźwięku)

Wysoka jakość dźwięku już przy przepływności 64 kb/s na kanał

Wiele kanałów (na potrzeby telewizji cyfrowej):

- do 48 głównych kanałów dźwiękowych
- do 16 kanałów niskotonowych
- do 16 kanałów wielojęzycznych
- do 16 kanałów danych

Lepsza jakość dźwięku przy lepszej kompresji

B

Kompresja MPEG 4 (dla dźwięku)

Format MPEG4 Audio jest to zbiór norm określających kodowanie naturalnego oraz syntetycznego dźwięku.

Części dotyczące kodowania syntetycznego to:

- SA (Structured Audio) oraz zwierają narzędzia umożliwiające symboliczną definicje muzyki, jak kiedyś. MIDI;
- mowy system TTSI (Text-to-Speech Interface),
- narzędzia do lokalizacji dźwięku 3D.

W części dotyczącwej kodowanie dźwięku naturalnego. Stosowane są 3 rodzaje kodeków:

- PARA (Parametric Codec) dla najniższych przepustowości,
- CELP (Code Excited Linear Predictive) dla średnich;
- TF (Time-Frequency), zawierający MPEG-2 AAC, który pozwala na uzyskanie najwyższej jakości dźwięku.

B

Kompresja MPEG 4 (dla dźwięku)

W formacie MPEG4 po raz pierwszy pojawia się pojęcie obiektu audio.

- Obiektem może być np. głos jednego lub kilku mówców, dźwięk jednego lub kilku instrumentów lub tło akustyczne. Obiekty audio mogą być grupowane oraz miksowane, przy czym rozdzielenie pojedynczego obiektu na mniejsze składowe jest już dosyć trudne.
- Każdy z obiektów dźwiękowych może być dowolnie modyfikowany, zależnie od potrzeb. W trudniejszych warunkach transmisji, możliwe jest ograniczenie pasma lub całkowite usunięcie obiektów nie mających wpływu na treść przekazu.

Format MPEG4 oferuje znacznie więcej niż jego poprzednicy: MPEG1 oraz MPEG2. Oprócz narzędzi umożliwiających kompresję i przesyłanie dźwięku naturalnego, wzbogacony został o możliwości syntezy dźwięku po stronie odbiorcy oraz skalowania przesyłanego strumienia bitów.



Kompresja MPEG 4 (dla dźwięku)

W formacie MPEG4 po raz pierwszy pojawia się pojęcie obiektu audio.

- Obiektem może być np. głos jednego lub kilku mówców, dźwięk jednego lub kilku instrumentów lub tło akustyczne. Obiekty audio mogą być grupowane oraz miksowane, przy czym rozdzielenie pojedynczego obiektu na mniejsze składowe jest już dosyć trudne.
- Każdy z obiektów dźwiękowych może być dowolnie modyfikowany, zależnie od potrzeb. W trudniejszych warunkach transmisji, możliwe jest ograniczenie pasma lub całkowite usunięcie obiektów nie mających wpływu na treść przekazu.

Format MPEG4 oferuje znacznie więcej niż jego poprzednicy: MPEG1 oraz MPEG2. Oprócz narzędzi umożliwiających kompresję i przesyłanie dźwięku naturalnego, wzbogacony został o możliwości syntezy dźwięku po stronie odbiorcy oraz skalowania przesyłanego strumienia bitów.

Kompresja MPEG 7



Trwają prace nad formatem MPEG-7, który ma połączyć większość aplikacji multimedialnych w jedną całość. Do przechowywania metadanych wykorzystuje XML

- ✓ Pozwoli to na łatwe zarządzanie multimedialnymi zasobami Internetu: wyszukiwanie, sprzedaż lub udostępnianie utworów oraz filmów.
- ✓ W zamierzeniu organizacji MPEG połączenie MPEG-4 i MPEG-7 (nazywane często MPEG-47) ma być optymalnym rozwiązaniem dla efektywnego przesyłu strumieniowego treści, jej zmiany i indeksowania.

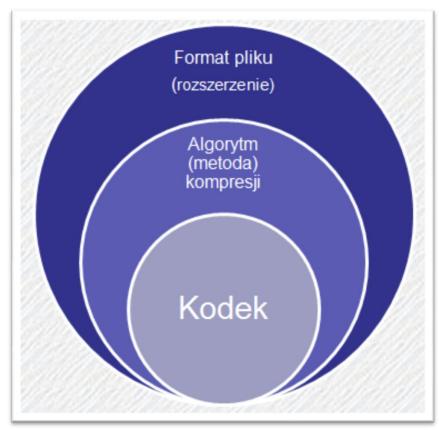
Kodek a format pliku

Kodek jest skrótem od "koder/dekoder", co oznacza urządzenie lub program zdolny do przekształcania strumienia

danych lub sygnału.

Przykład:

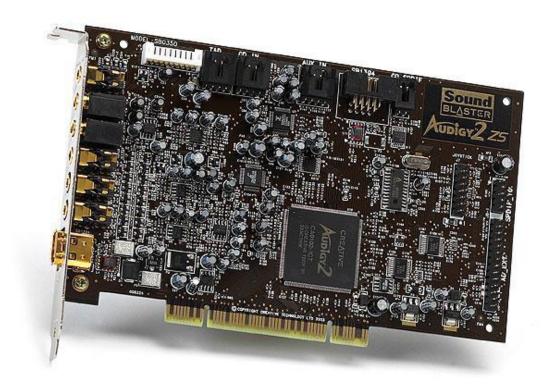
Plik .avi oprócz obrazu zawiera także dźwięk. Dźwięk ten może być skompresowany metodą MPEG leyer III (mp3) przy użyciu kodeka LAME



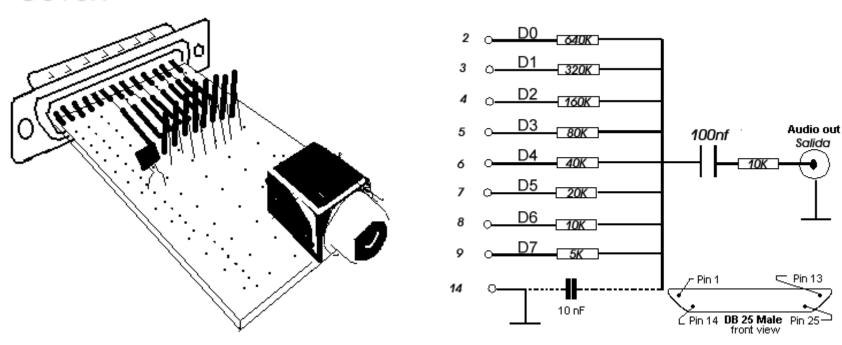
Część 11



Karta dźwiękowa



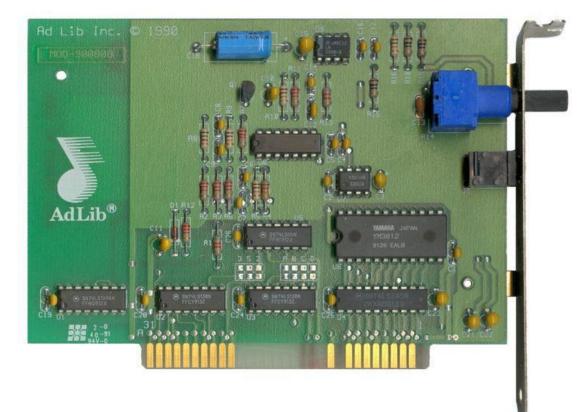
Covox



Wprowadzony na rynek w 1986 r. Prosty zewnętrzny przetwornik audio, dołączany do portu drukarki komputera by wyprowadzić dźwięk. Zbudowany z najprostszego 8-bitowego przetwornika cyfrowoanalogowego opartego na drabince rezystorowej.



AdLib Music Synthesizer Card 1987 r.



Pierwsza masowo sprzedawana karta dźwiękowa dla komputerów PC

Wykorzystywała syntezę FM





GravisUltrasound 1993 r.



wyposażona w syntezator PCM, jednak jej głównym atutem była Synteza Wavetable

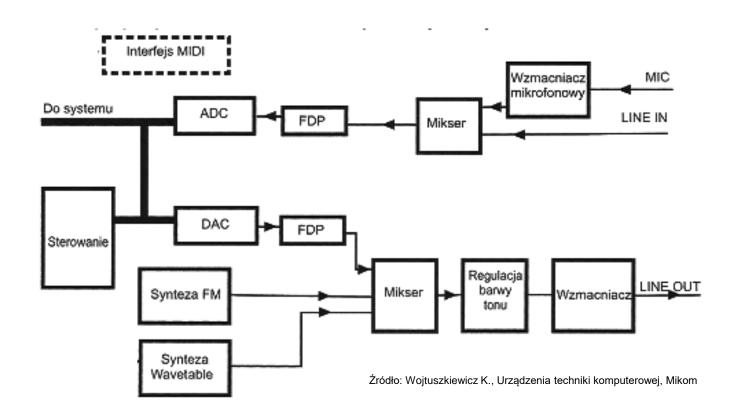
- Odtwarzanie 16-bit 44.1kHz
- Nagrywanie 8-bit, 44 kHz dr Artur Bartoszewski Informatyka i systemy informatyczne , sem. 1- WYKŁAD



Sound Blaster 1992 r.

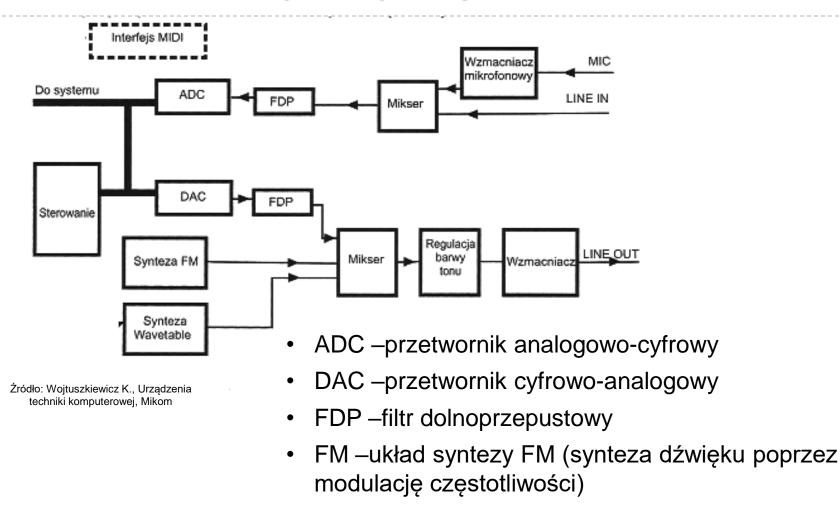


- generator FM OPL3 (zgodny z Yamaha YM3812), przetwornik 8-bit / 44,1 kHz stereo,
- port Game/MPU401,
- ISA 16-bit



B

Zasad działania karty dźwiękowej



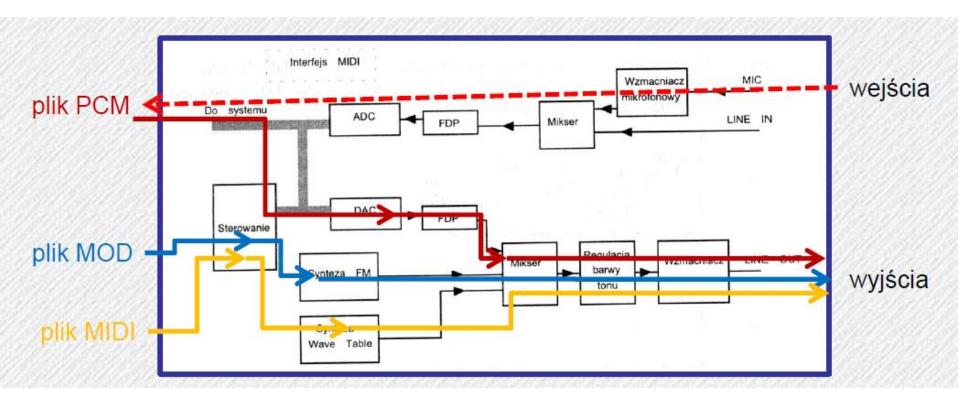
MIDI

Źródło: Wojtuszkiewicz K., Urządzenia techniki komputerowej, Mikom

WaveTable –tablica próbek dźwięku dla syntezy



Zasad działania karty dźwiękowej



Przetworniki AD/DA



Przetworniki AD/DA (ang. analog-digital / digital analog) (pol. AC/CA – analogowo-cyfrowe / cyfrowo-analogowe) układy sprzętowe

 AD - wykonujące digitalizację - zamieniające sygnał audio na zapis cyfrowy PCM,



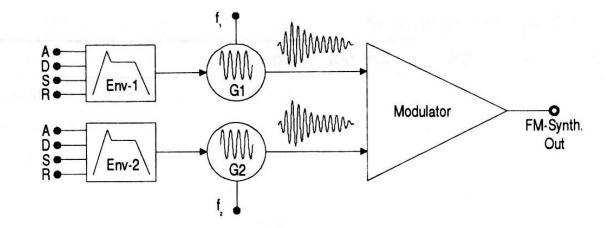
 DA - generujące sygnał elektryczny audio na podstawie cyfrowego zapisu PCM





Synteza FM pozwala na generowanie programowe dźwięków.

Polega na wykorzystaniu interferencji dwóch przebiegów o nieznacznie różniących się częstotliwościach.



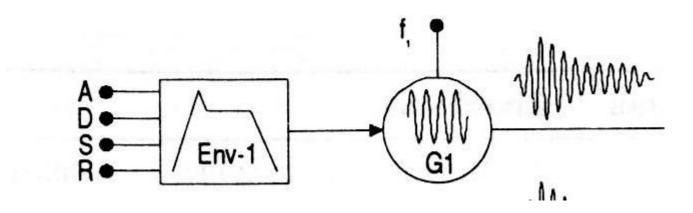
Modulacja FM: $FM(t) = A_C \times \sin(2\pi f_C \times t + A_M \times \sin(2\pi f_M \times t))$

Sumowanie: $SUM(t) = A_C \times \sin(2\pi f_C \times t) + A_M \times \sin(2\pi f_M \times t)$



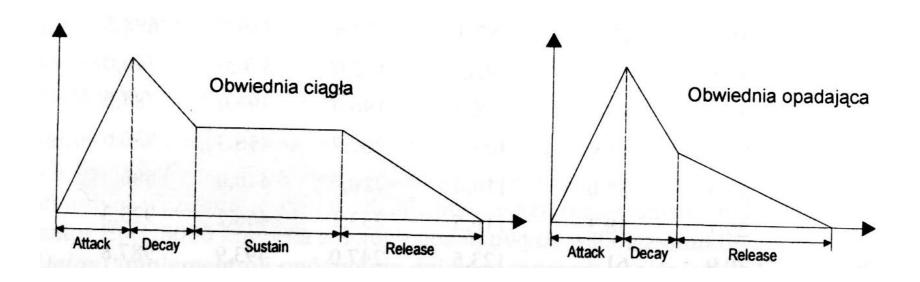
Podstawowym składnikiem syntezatora FM jest operator zbudowany z:

- oscylatora,
- generatora obwiedni,
- generatora efektów,
- sterownika głośności,
- filtru dolnoprzepustowego.



Źródło: Metzger P., Anatomia PC

Aby generowany dźwięk był zbliżony do naturalnego generowanemu przebiegowi nadawana jest obwiednia (wzorowana na naturalnych źródłach dźwięku)



Źródło: Metzger P., Anatomia PC

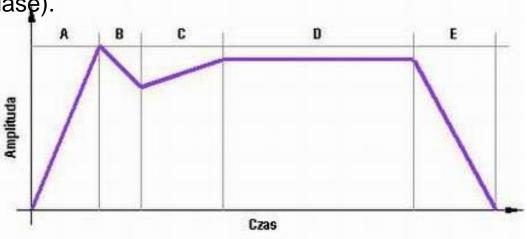


Amplituda dźwięku wytwarzanego przez instrument podlega pewnym zmianom w czasie: początkowo narasta (najczęściej bardzo szybko), potem pozostaje w pobliżu pewnej ustalonej wartości pośredniej i dopiero wtedy opada do zera.

Te zmiany amplitudy nazywamy obwiednią i można je podzielić na pięć faz:

- A. faza nabrzmienia (ang. attack),
- B. faza opadania (ang. decay),
- C. stadium dochodzenia do ustalonej fazy (ang. slope),
- D. faza ustalona (ang. sustain),

E. zanikanie dźwięku (ang. relase).

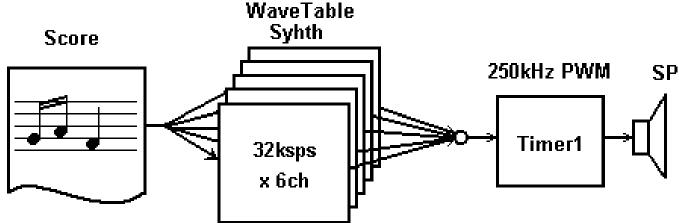


Synteza Wavetable

Synteza Wavetable - z uwagi na sztuczne brzmienie generowanych dźwięków synteza FM nie nadaję się do zastosowań profesjonalnych. Z tego powodu opracowano technikę syntezy wavetable (WT).

W celu uzyskania brzmienia instrumentu chip muzyczny nie generuje sztucznego dźwięku, lecz odtwarza oryginalny dźwięk instrumentu, nagrany wcześniej w studiu.

W praktyce nie ma możliwości zapisania w pamięci wszystkich możliwych dźwięków generowanych przez 128 instrumentów MIDI. Chip muzyczny musi więc ok " róbek.



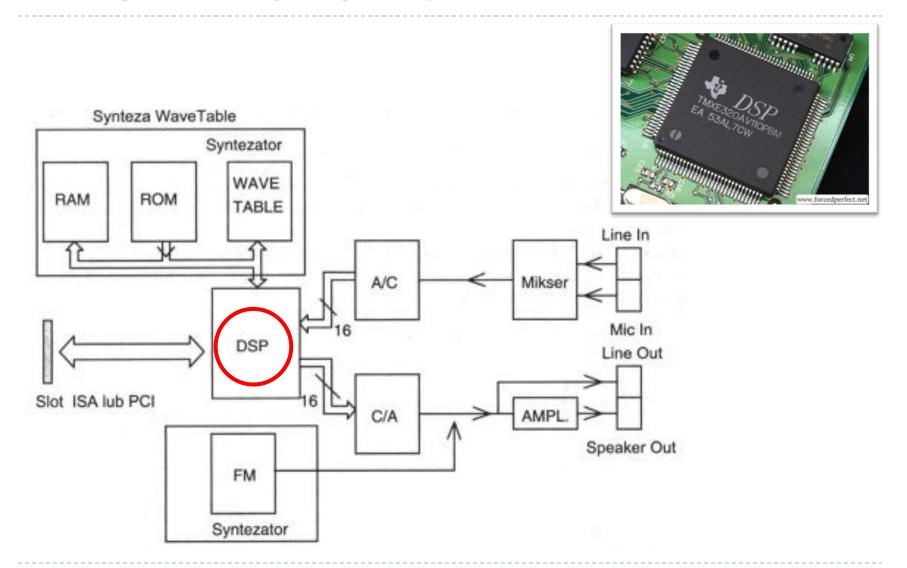
Synteza Wavetable

Synteza Wawetable opiera się na koncepcji MIDI cyfrowe złącze instrumentów muzycznych (MIDI), wprowadzone na rynek we wczesnych latach 80.

- Komunikaty MIDI mogą być proste (np. włącz dźwięk pianina na 5 sekund), lub złożone (np. zwiększyć napięcie wzmacniacza VCA w generatorze 6, aby dopasować częstotliwość do generatora nr1).
- Należy pamiętać, że MIDI nie przesyła dźwięku lecz informacje o nim.



Rozwój kart dźwiękowych – procesor DSP





Rozwój kart dźwiękowych – procesor DSP

Procesor DSP (ang. Digital Signal Processor) - Procesor sygnałowy – klasa procesorów specjalizowanych do cyfrowej obróbki sygnałów.

Posiadaniem osobne pamięci przeznaczone dla programu i dla danych (architektura harwardzka)



Procesor DSP jest właściwie specjalizowanym mikrokomputerem jednoukładowym

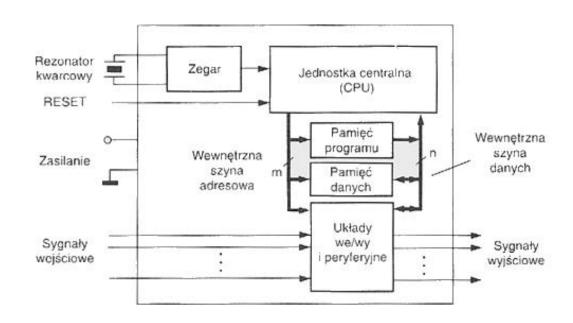


Procesor DSP



Główne elementy typowego procesora DSP to:

- rdzeń procesora (CPU)
- pamięci z podziałem na pamięć danych i pamięć programu
- koprocesor DMA
- obwody współpracy ze środowiskiem zewnętrznym



Procesor DSP



DSP znalazły zastosowanie w:

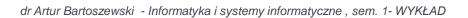
- odtwarzaczach MP3, CD itp.
- telefonach komórkowych
- aparatach cyfrowych kamerach
- kontrolerach napędów dysków
- modemach i routerach



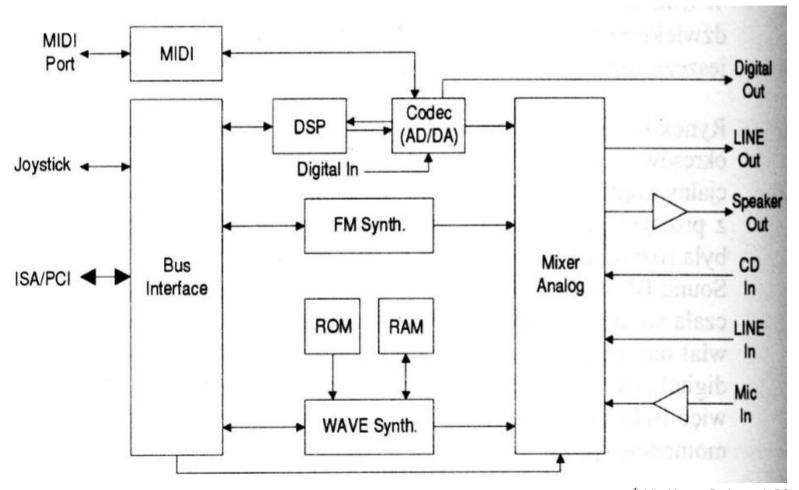








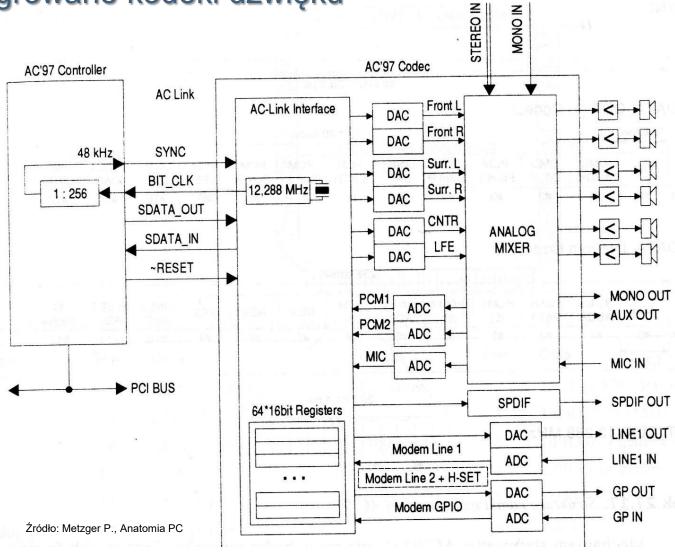
Schemat karty dźwiękowej - II



Źródło: Metzger P., Anatomia PC



Zintegrowane kodeki dźwięku



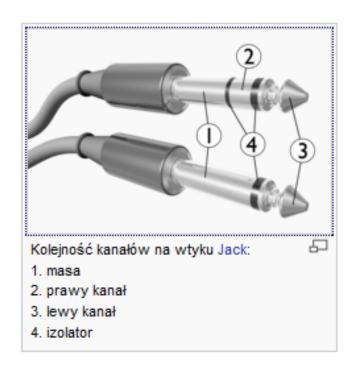
Złącza karty dźwiękowej

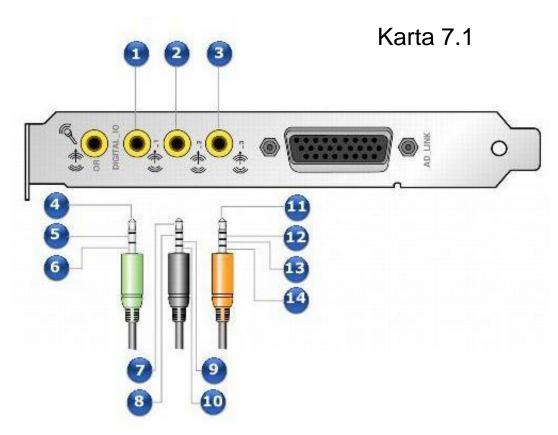


Kolor	Funkcja
różowy	Analogowe wejście dla mikrofonu.
błękitny	Analogowe wejście audio. (line-in)
jasnozielony	Analogowe wyjście dla głośników albo słuchawek, w systemach wielogłośnikowych wyjście dla przednich głośników
czarny	Analogowe wyjście dla głośników tylnych.
pomarańczowy	Cyfrowe wyjście dźwięku (S/PDIF), czasami tym kolorem oznacza się analogowe wyjście dla głośników centralnego i niskotonowego.

Złącza karty dźwiękowej



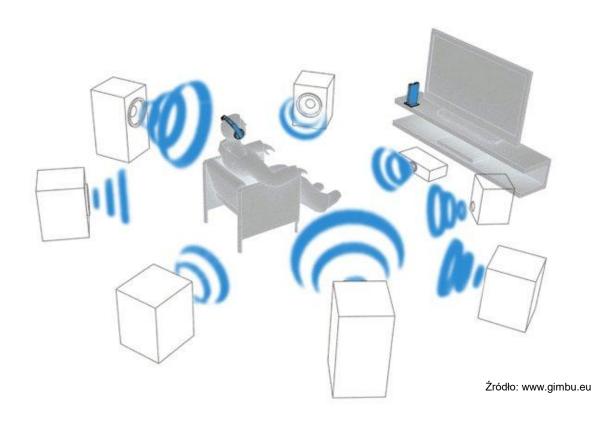




Część 1II

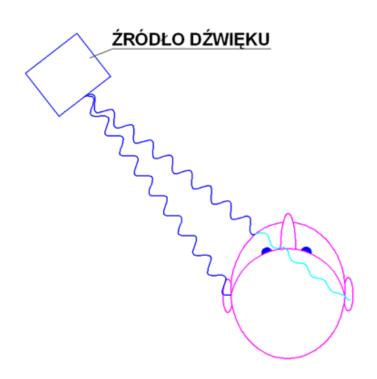


Dźwięk przestrzenny



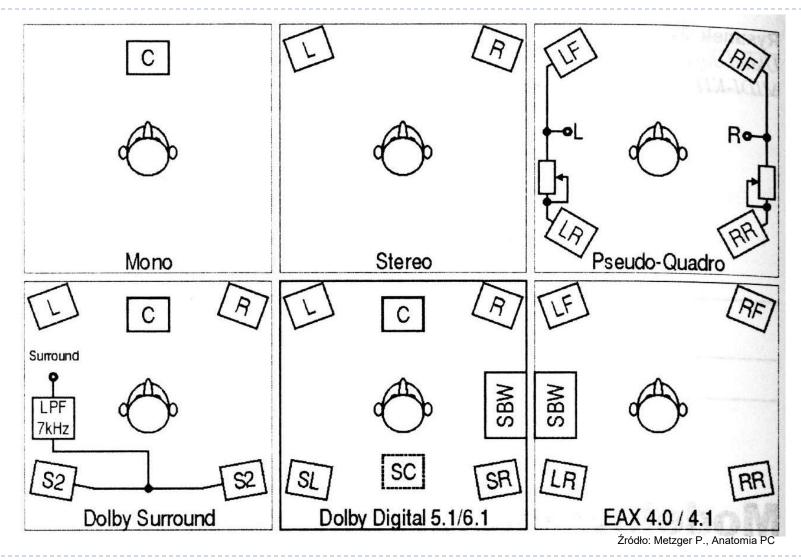
Pozycjonowanie dźwięku





- ✓ Dźwięk pozycjonujemy (rozpoznajemy skąd dochodzi) dzięki przesunięciu w czasie i różnicy w głośności i barwie dźwięku dochodzącego do obu uszu.
- ✓ Dobrze pozycjonujemy tony wysokie i średnie, gorzej albo wcale tony niskie.

Systemy dźwięku przestrzennego



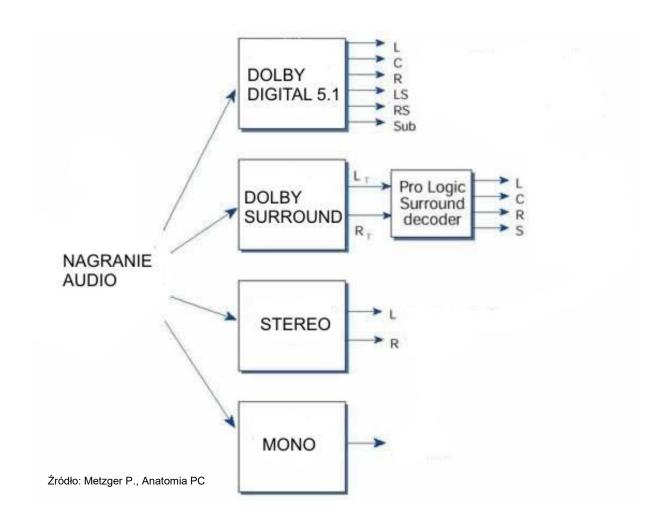
Systemy dźwięku przestrzennego

- ✓ Mono
- ✓ Stereo –d wukanałowy
- ✓ Dolby Surround format dźwięku pseudoprzestrzennego otrzymywany z formatu dwukanałowego (obecnie rzadko stosowany); po zdekodowaniu źródła stereofonicznego otrzymujemy 3 kanały (lewy, prawy i surround)
- ✓ Dolby Surround Pro Logic format kodowania dźwięku otaczającego, dookólnego (ang. surround-otaczać). Jest to ulepszona wersja systemu dźwieku przestrzennego (stereo) z czterema kanałami: przednim lewym, przednim centralnym i przednim prawym oraz kanałem tylnym. System został opracowany przez Dolby Laboratories i wprowadzony na rynek w 1987 roku. Był to pierwszy system dźwięku dookólnego surround powszechnie przyjęty do użytku domowego.
- ✓ Dolby Digital system cyfrowego kodowania dźwięku sześciokanałowego o maksymalnej częstotliwości próbkowania 96 kHz, 24-bitowej rozdzielczości (standard sygnału audio zapisywanego na DVD –mimo to w większości filmów stosuje się: 16-bit, 48 kHzlub 44,1 kHz), o dynamice 105 dB, kompresji 1:12 i wymaganej przepustowości 448 kb/s.
- ✓ DolbyDigital 5.1to 5 szerokopasmowych kanałów (od 20 do 20000 Hz) przeznaczonych dla 5 głośników (lewego, centralnego, prawego, lewego surround, prawego surround); szósty kanał jest przeznaczony dla głośnika niskotonowego (sub-

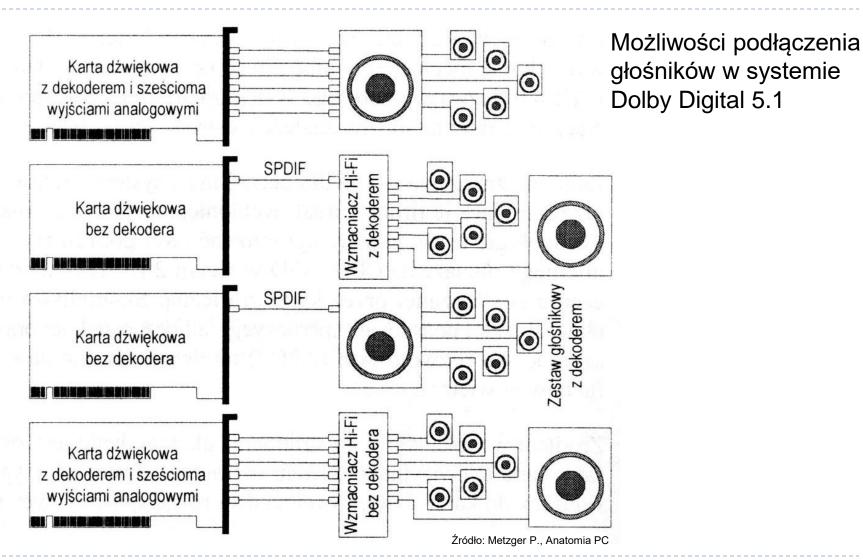
55woofera–20-120 Hz).



Systemy dźwięku przestrzennego



Systemy dźwięku przestrzennego





Literatura:

W prezentacji wykorzystano fragmenty i zadania z książek i prezentacji:

- Wojtuszkiewicz K., Urządzenia techniki komputerowej, Mikom
- Nasiłowski D., Jakościowe aspekty kompresji obrazu i dźwięku, MIKOM, Warszawa 2004,
- Metzger, p. Anatomia PC Wydanie XI, Helion 2007,
- http://www.inf.sgsp.edu.pl/
- http://deuter.am.put.poznan.pl/zwm/