1. **Klasyfikacja metod syntezy dźwięku**

W zależności od liczby możliwych do jednoczesnego wygenerowania dźwięków wyróżniamy 2 typy syntezy:

* Monofoniczna (homofoniczna) - w starszych syntezatorach analogowych, lub przy dużej złożoności obliczeniowej syntezy - w danej chwili można wygenerować tylko pojedynczy dźwięk
* Polifoniczna - w danej chwili można wygenerować jednocześnie wiele dźwięków

Odrębną kategorię stanowią syntezatory wielobrzmieniowe, gdzie można jednocześnie wygenerować dźwięki o różnych barwach.

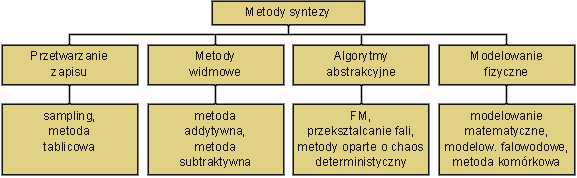
Głos w syntezatorach jest torem syntezy (w odróżnieniu od głosu organowego, gdzie oznacza brzmienie, zbiór piszczałek).

Jednoczesna synteza dźwięków o różnych barwach może odbywać się poprzez:

* Layering - uwarstwianie brzmień (dla dźwięków o tej samej wysokości)
* Splitting - podział klawiatury na strefy

Klasyfikację syntezatorów można również przeprowadzić w zależności od metody syntezy na:

* Syntezatory analogowe
* Syntezatory cyfrowe

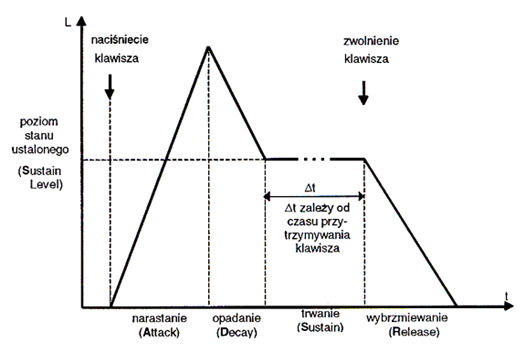


1. **Obwiednia ADSR**

W syntezatorach zawarty jest także zwykle generator obwiedni ADSR, zmieniający głośność dźwięku w czasie jego trwania i tym samym upodabniający dźwięk do dźwięków zwykłych instrumentów muzycznych.

Ponadto w syntezatorach stosowany jest często generator wolnych przebiegów w celu uzyskania takich efektów, jak wibrato czy tremolo.

Model obwiedni ADSR przedstawia poniższy rysunek.



Faza stanu ustalonego zależy tu od grającego.

1. **Na czym polega synteza addytywna i subtraktywna dźwięku?**

Synteza dźwięku w metodach widmowych polega na zastosowaniu odpowiednich operacji na widmie dźwięku.

Do metod widmowych zaliczamy następujące metody syntezy dźwięku:

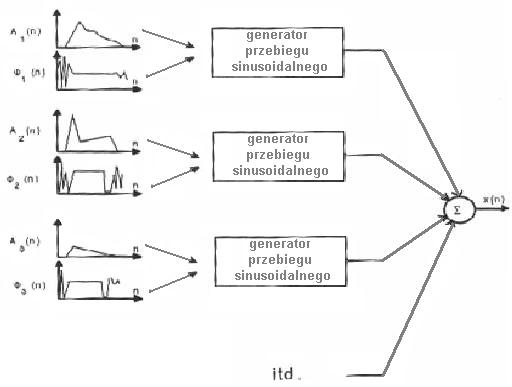
* Metoda addytywna
* Metoda subtraktywna

W metodzie addytywnej widmo dźwięku "budowane" jest z pojedynczych składowych harmonicznych, z których każda może być modulowana amplitudowo i fazowo.

Aby uzyskać brzmienie symulujące brzmienie instrumentów, dźwięki instrumentów akustycznych są poddawane analizie widmowej (FFT), na podstawie której przeprowadzana jest resynteza.

Metoda ta jest rzadko stosowana w elektronicznych instrumentach muzycznych.

Schemat syntezatora addytywnego przedstawia poniższy rysunek.



Metoda addytywna ma różne odmiany:

* PV (Phase Vocoder) - dla sygnału okresowego
* MQ (McAulay-Quatieri) - dla dźwięków nieharmonicznych

Metoda subtraktywna polega na odejmowaniu określonych składowych widma z szumu lub sygnału szerokopasmowego w układzie filtracyjnym.

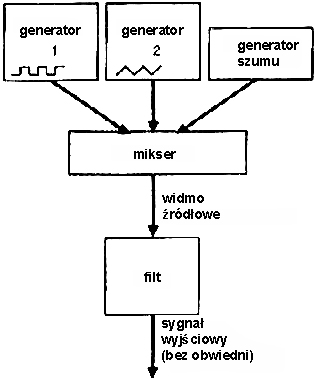
Proces filtracji w modelu ARMA (AR+MA: autoregression - moving average) opisywany jest poniższym wzorem:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w6/Image4330.gif

gdzie http://edu.pjwstk.edu.pl/wyklady/mul/scb/w6/Image4331.gif - współczynniki filtru,  
u - wymuszenie,  
G - wzmocnienie filtru.

Metoda subtraktywna stosowana jest zarówno w syntezatorach analogowych, jak i cyfrowych.

Ogólny schemat tej metody przedstawia poniższy rysunek.



1. **Na czym polega synteza dźwięku metodą modelowania fizycznego? Wymienić rodzaje tej syntezy**

Syntezatory działające w oparciu o modele fizyczne instrumentów akustycznych symulują zjawiska fizyczne zachodzące w tych instrumentach, przy uwzględnieniu modelowania procesów artykulacyjnych.

Takie modelowanie jest szczególnie przydatne do syntezy gitary, saksofonu, trąbki, fletu, czy piszczałek organowych.

Do grupy modelowania fizycznego zaliczamy następujące metody syntezy dźwięku:

* Synteza komórkowa
* Modelowanie matematyczne
* Modelowanie falowodowe

1. **Na czym polega synteza dźwięku metodą modelowania matematycznego oraz falowodowego? Podać postać równania struny nieskończonej oraz równania fali płaskiej w nieskończonym cylindrze**

Metoda modelowania matematycznego polega na matematycznym modelowaniu zjawisk falowych.

Rozwiązuje się równanie falowe, opisujące drgania w danym ośrodku (struny, słupa powietrza itp.), a funkcja będąca rozwiązaniem równania stanowi przebieg czasowy dźwięku syntetycznego.

Zaleta metody modelowania matematycznego jest możliwość rzetelnego sprawdzenia wiarygodności modeli opisujących rzeczywiste instrumenty.

Wada tej metody jest konieczność stosowania złożonych obliczeniowo algorytmów numerycznych, pozwalających na całkowanie równania falowego.

Równania różniczkowe zwykle zapisuje się w postaci równań różnicowych, ale złożoność obliczeniowa jest nadal duża.

Metody modelowania matematycznego są metodami eksperymentalnymi i nie są stosowane w syntezatorowych instrumentach muzycznych.

Modelowanie falowodowe (ang. Digital Waveguide Modeling) polega na modelowaniu przy pomocy cyfrowego falowodu propagacji fal bieżących, składających się na falę stojącą w danym instrumencie.

Zaletą tej metody jest stosunkowo mała złożoność obliczeniowa i możliwość syntezy w czasie rzeczywistym.

Metoda modelowania falowodowego znalazła zastosowanie w syntezie:

* Instrumentów muzycznych - opracowano falowodowe modele gitary, klarnetu, fortepianu, fletu itd.
* Syntezie mowy i śpiewu

Szczegóły implementacji modelowania falowodowego zostaną omówione dla:

* Modelowania instrumentów strunowych
* Modelowania instrumentów dętych

Równanie opisujące przemieszczenie struny ma następującą postać:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4362.gif, gdzie:

y - wychylenie struny

K - naprężenie struny

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4363.gif- liniowa gęstość masy

x - współrzędna leżąca na osi równoległej do struny

Rozwiązaniem tego równania jest złożenie 2 fal bieżących, przemieszczających się w prawo i w lewo z prędkością c

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4364.gif

Rozwiązanie to można zapisać w postaci

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4365.gif

Zmienne ciągłe zastępowane są dyskretnymi (zakłada się, ze fale bieżące mają ograniczone widmo)

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4366.gif

gdzie:

T - okres próbkowania,  
X - przestrzenny interwał próbkowania.

Definiując

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4367.gif

rozwiązanie równania struny można zapisać w postaci:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4368.gif

Wyrażenie http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4369.gif można interpretować jako wartość próbki na wyjściu linii opóźniającej o m próbek, do wejścia której doprowadzany jest sygnał http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4370.gif, zaś wyrażenie http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4371.gif jako wejście linii opóźniającej o m próbek.

Uwzględnienie strat, spowodowanych m.in. tarciem wewnętrznym i oporem powietrza daje rozbudowę modelu do postaci:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4375.gif

gdzie http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4376.gif - współczynnik proporcjonalności strat.

Rozwiązanie ogólne tego równania ma postać

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4377.gif

Fale bieżące są więc tłumione eksponencjalnie.

Po spróbkowaniu rozwiązanie ma postać:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4378.gif

gdzie http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4379.gif

Jednowymiarowe równanie falowe fali płaskiej wewnątrz nieskończenie długiego cylindra ma postać:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4398.gif

gdzie:

c - prędkość propagacji dźwięku w powietrzu (ok. 340 m/s),  
p - ciśnienie,  
x - odległość wzdłuż osi korpusu piszczałki.

Rozwiązanie ogólne równania falowego ma postać:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4399.gif, gdzie:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4400.gif - dowolne, podwójnie różniczkowalne funkcje ciągłe,  
http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4401.gif- fala bieżąca ciśnienia przemieszczająca się w kierunku dodatnich wartości x z prędkością c,  
http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4402.gif - fala w kierunku ujemnych x, przemieszczająca się z prędkością c.

Uproszczone modelowanie bezstratnego cylindra o ograniczonej długości odbywa się za pomocą filtrów:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4403.gif - modelowanie odbicia fal o częstotliwościach poniżej pewnej częstotliwości granicznej,  
http://edu.pjwstk.edu.pl/wyklady/mul/scb/w7/Image4404.gif - modelowanie wypromieniowania z cylindra fal o częstotliwościach powyżej pewnej częstotliwości granicznej.

Częstotliwość graniczna jest funkcją średnicy cylindra.

1. **Wzrok i słuch - podstawowe pojęcia: bodziec wzrokowy/słuchowy, wrażenie, percepcja**

Bodziec wzrokowy/słuchowy to mierzalna cecha fizyczna, np. luminancja, długość fali, ciśnienie (w przypadku dźwięku) etc.

Wrażenie to natychmiastowy efekt pobudzenia fizycznego.

Percepcja to efekt zjawiska zmysłowego; na efekt ten wpływają procesy wyższego rzędu, takie, jak pamięć, uwaga, czy doświadczenie.

1. **Składowe przestrzeni koloru w terminach obserwatora i w terminach światła**

Barwa jest psychofizyczną cechą percepcji wzrokowej - jest wrażeniem wytwarzanym w mózgu jako reakcja na światło dopływające do siatkówki oka. Wrażenie barwy jest zależne od rodzaju światła emitowanego przez źródła światła lub odbijanego przez obiekty.

Barwa jest wrażeniem subiektywnym, zależnym od oceny obserwatora, który odbiera (i przetwarza) bodziec. Wiele różnych widm jest postrzeganych jako ta sama barwa; zjawisko to określamy mianem metameryzmu.

Barwa jest opisywana w przestrzeni trójwymiarowej.

Atrybuty barwy postrzeganej w terminach obserwatora są następujące:

* kolor
* jasność/intensywność/luminancja
* nasycenie/chrominancja

Kolor jest atrybutem określającym "kolor" widziany przez człowieka (np. żółty, czerwony, itp.).

Jasność/intensywność/luminancja opisują ilość światła emitowanego lub odbijanego - "ile czerni jest wmieszane w kolor".

Luminancja jest achromatycznym składnikiem barwy, określającym ilość światła emitowanego lub odbitego.

Intensywność dotyczy barw achromatycznych. Jest wielkością fizyczną (cd/m2), ale czasem używana jest zamiennie z terminami perceptualnymi, psychofizycznymi.

Jasność opisuje ilość światła odbijanego przez obiekty lub emitowanego przez źródła światła.

W odniesieniu do obiektów i ilości odbijanego światła stosuje się następującą skalę w opisach słownych: bardzo jasny, jasny, średni, ciemny, bardzo ciemny.

W odniesieniu do źródeł światła i ilości światła emitowanego stosuje się następującą skalę w opisach słownych: bardzo przyćmiony, przyćmiony, średni, jasny, bardzo jasny.

Nasycenie/chrominancja określa czystość (w terminach zmieszania z bielą) lub żywość barwy. Jest to stopień zróżnicowania w stosunku do szarości dla różnych barw o tej samej jasności.

Nasycenie dotyczy jasności barwy - zwiększenie jasności jest postrzegane jako zmniejszenie nasycenia i na odwrót.

Chrominancja jest określana przez porównanie z bielą i nie zmienia się przy zmianie jasności. W odniesieniu do chrominancji stosuje się następujące opisy słowne: szarawa, średnia, mocna, żywa.

Barwę można również opisywać w terminach światła.

Atrybuty barwy specyfikowanej w terminach światła są następujące:

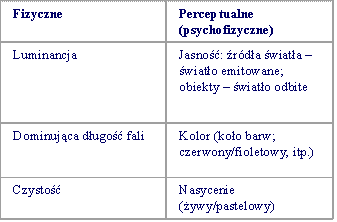
* dominująca długość fali
* luminancja
* czystość

Dominująca długość fali określa kolor - odpowiada subiektywnemu wrażeniu koloru.

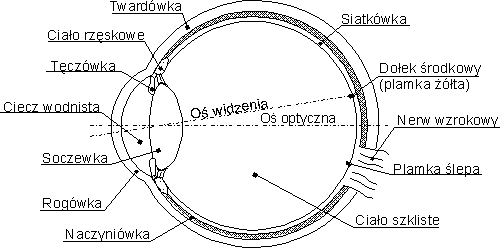
Luminancja określa ilość światła emitowanego lub światła odbitego. Dla światła achromatycznego jest to intensywność światła, dla chromatycznego odpowiada subiektywnemu pojęciu jasności.

Czystość określa rozkład widmowy, jaki powoduje powstanie światła o danej barwie. Jest to proporcja dominującej długości fali i światła białego, niezbędnego do zdefiniowania tej barwy. Czystość odpowiada perceptualnemu pojęciu nasycenia.

Istnieje wzajemna odpowiedniość między fizycznymi (opartymi na świetle) i perceptualnymi (postrzeganymi przez obserwatora) terminami opisu barwy, przedstawiona w tabelce poniżej.



1. **Budowa oka**



1. **Czułość wzroku i rozdzielczość wzroku**

Czułość wzroku jest to zdolność ekstrakcji informacji przy niskich poziomach luminancji (stopień reakcji na światło).

Rozdzielczość wzroku jest to zdolność rozróżniania małych detali przestrzennych.

1. **Addytywne/subtraktywne mieszanie kolorów i kolory podstawowe**

Mieszanie addytywne jest procesem mieszania emisji ze źródeł światła, pokrywających różne części widma. Czerń otrzymywana jest, gdy nie miesza się żadnych barw (zerowa energia emisji), zaś biel otrzymywana jest jako zmieszanie 3 barw podstawowych z maksymalną energią. Przykładem zastosowania mieszania addytywnego jest telewizja kolorowa, gdzie stosowane są barwy podstawowe RGB (Red-Green-Blue), tj. czerwona, zielona i niebieska.

Mieszanie subtraktywne to proces filtracji odbijanych części widma; mieszanie farb. W tym procesie rozkład widmowy światła ulega selektywnej redukcji. Użyta w mieszaniu farba usuwa ze światła białego te składowe, które pochłonęła, a odbija fale odpowiadające jej barwie. Biel powstaje przy braku mieszania (odbijanie całego widma; nieprzezroczysta biała powierzchnia odbija i rozprasza wszystkie długości fal). Czerń otrzymuje się poprzez zmieszanie 3 barw podstawowych w maksymalnych ilościach (czarna powierzchnia pochłania wszystkie długości fal). Przykładem zastosowania jest druk kolorowy, gdzie używany jest model CMY (Cyan-Magenta-Yellow, tj. zielonkawoniebieski-purpurowy-żółty), oraz jego rozszerzona wersja CMYK(CMY-Black, tj. dodany barwnik czarny).

Barwy CMY stosowane w druku w mieszaniu subtraktywnym są odpowiednio dopełnieniami barw, używanych do opisu RGB do światła białego. Dzięki tej właściwości CMYK lepiej nadaje się do opisu barwy w druku, gdzie rolę gra pochłanianie światła przez farbę drukarską.

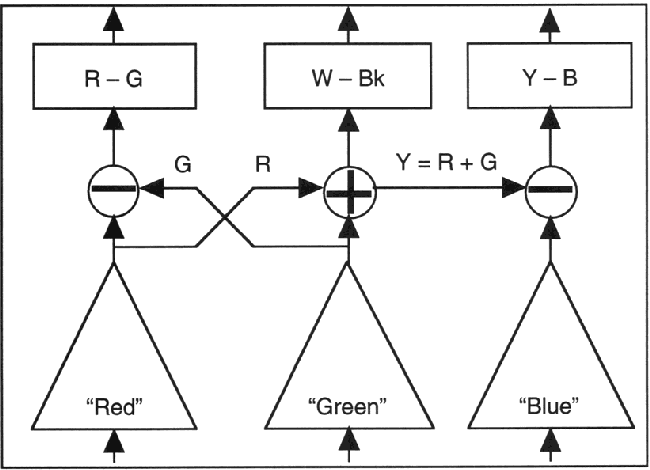
1. **Kanały kolorów w nerwie optycznym**

Sposób widzenia barwy przez człowieka związany jest ze sposobem przekazywania impulsów nerwowych w układzie wzrokowym.

Wyjścia fotoreceptorów oka łączą się w nerwie optycznym, gdzie są konwertowane do następujących 3 kanałów:

* R+G
* R-G
* Y-B (Y - yellow, żółty)

Ilustruje to poniższy rysunek.



(źródło: Levkowitz, H., Color Theory and Modeling for Computer Graphics, Visualisation, and Multimedia Applications, Kluwer AP, Boston, USA, 1997)

Kombinacje barw w tych kanałach są następujące:

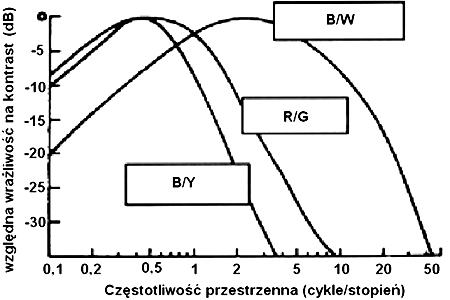
* R+G: achromatyczna zawartość barwy (jasność). Błękit jest z tego kanału wykluczany - nie wpływa na percepcję jasności; zmiany błękitu nie są wystarczające do dostrzeżenia zmian barwy, a zatem nie jest konieczne ich kodowanie. Kanał ten jest opisywany inaczej jako W-Bk (white minus black - biały minus czarny)
* R-G: jeden z kanałów chromatycznych ("czerwony-lub-zielony", "czerwony-minus-zielony"). Stąd nie jest możliwe uzyskanie wrażenia koloru czerwono-zielonego, w odróżnieniu od możliwości zobaczenia koloru zielono-żółtego, zielono-niebieskiego, czerwono-niebieskiego)
* Y-B: drugi z kanałów chromatycznych ("żółty-lub-niebieski", "żółty-minus-niebieski"). Stąd nie jest możliwe uzyskanie wrażenia koloru żółto-niebieskiego

1. **Związek luminancji z rozdzielczością widzenia barw**

Mechanizmy percepcji koloru i luminancji uzupełniają się nawzajem. Interakcje te szczególnie dotyczą rozdzielczości wzroku.

Luminancja ma wpływ przede wszystkim na postrzeganą rozdzielczość.

System percepcji luminancji może rozróżniać bardzo drobne szczegóły. Szczyt czułości przypada na 2-4 cykle na stopień (patrz rysunek poniżej), z częstotliwością odcięcia ok. 60 cykli na stopień.



(źródło: Levkowitz, H., Color Theory and Modeling for Computer Graphics, Visualisation, and Multimedia Applications, Kluwer AP, Boston, USA, 1997)

System percepcji koloru może rozróżnić jedynie zgrubne zmiany przestrzenne. Szczyt czułości dla krat o stałej luminancji przypada dla niskich częstotliwości, a częstotliwość odcięcia - około 10-20 cykli na stopień.

Postrzeganie luminancji jest związany z postrzeganą rozdzielczością barw w następujący sposób:

* Mechanizm postrzegania luminancji pośredniczy w postrzeganiu wysokiej częstotliwości przestrzennej
* Mechanizm luminancji ma szersze pasmo
* Mechanizm widzenia kolorów pośredniczy w postrzeganiu niskiej częstotliwości przestrzennej

Mechanizm luminancji pośredniczy w postrzeganiu wysokiej częstotliwości przestrzennej, a zatem np. trudno odczytać żółty tekst na białym tle, gdyż różnica w luminancji jest wówczas bardzo mała.

Wysoka rozdzielczość przestrzenna zależy od luminancji i jest niezależna od koloru.

Mechanizm luminancji ma szersze pasmo, a zatem konieczne jest szersze pasmo do zakodowania przestrzennych zmian luminancji. Dodanie barwy przy konwersji telewizji czarno-białej do kolorowej wymagało niewiele dodatkowego pasma.

W kompresji obrazów barwnych należy poświęcić większe pasmo na kodowanie luminancji, aby osiągnąć wyższy stosunek kompresji nie tracąc na jakości obrazu

Mechanizm widzenia kolorów pośredniczy w postrzeganiu niskiej częstotliwości przestrzennej, stąd małe punkty kolorowe "tracą" kolor i wydają się być achromatyczne. Barwy wydają się być bardziej nasycone i intensywne na większych powierzchniach. Obraz na wielkim ekranie wydaje się być bardziej kolorowy niż na mniejszym.

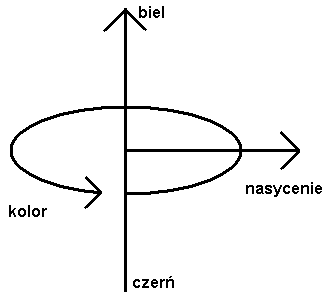
Przy projektowaniu GUI (graphical user interfaces) małe różnice koloru między oknami są wystarczające do odróżnienia okien i nie ma potrzeby używania ostrych, nasyconych barw.

1. **Wymienić przykłady modeli przestrzeni barw**

Przestrzeń barw jest trójwymiarowa.

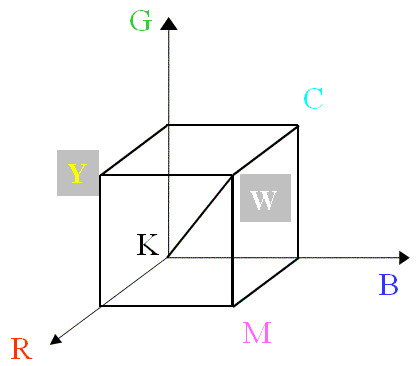
Istnieje wiele modeli przestrzeni barw, np. HSB/HSL (kolor, nasycenie, jasność), czy RGB.

Poniższy rysunek przedstawia model HSB/HSL. Model ten odpowiada sposobowi określania barwy stosowanemu przez artystów.



(źródło: Levkowitz, H., Color Theory and Modeling for Computer Graphics, Visualisation, and Multimedia Applications, Kluwer AP, Boston, USA, 1997)

Model RGB przedstawiono poniżej.



Stosowane są różne systemy barw, np.:

* Systemy fizyczne: RGB, CMY(K)
* Systemy kolorymetryczne: CIE XYZ, opracowany przez CIE (Commision Internationale de l'Eclairage)
* Systemy perceptualne: CIE La\*b\*, CIELUV, HSB (Munsella), HLS
* Systemy naturalne (nazewnicze): czerwony, zielony, niebieski; jasny, średni, ciemny

Zakres barwowy otrzymywany w różnych systemach i na różnych urządzeniach jest zróżnicowany.

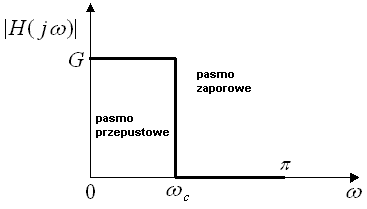
Gamut to zakres barwowy, który określa możliwości odwzorowania barw i dotyczy wszystkich urządzeń technologicznych tworzących barwy.

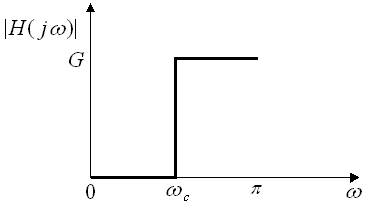
1. **Podziały filtrów, charakterystyki amplitudowe filtrów**

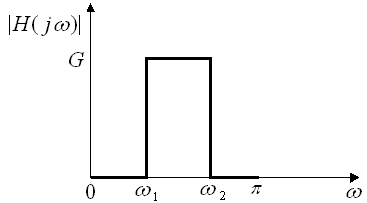
W przetwarzaniu sygnałów stosowane są następujące filtry:

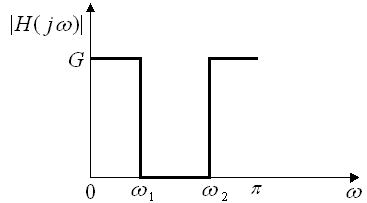
* Dolnoprzepustowy
* Górnoprzepustowy
* Pasmowoprzepustowy
* Pasmowozaporowy
* Grzebieniowy
* Wszechprzepustowy (korektor fazy)

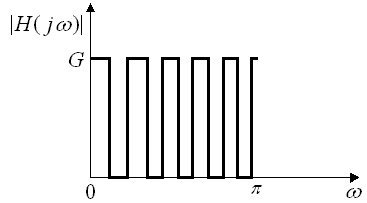
Charakterystyki filtrów idealnych ww. typów przedstawiono na poniższych rysunkach, na przykładzie filtru o wzmocnieniu G, z zachowaniem powyższej kolejności (źródło: <http://server.eletel.p.lodz.pl/~materka/ps5.pdf>).

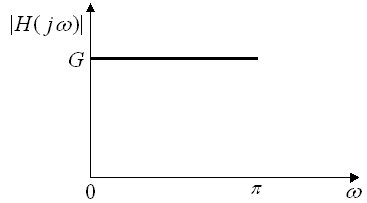












Filtry opisuje się również poprzez pasmo trzydecybelowe, w którym pasmo przenoszenia spada o 3 dB.

Pasmo 3dB jest pasmem pomiędzy 2 granicznymi wartościami częstotliwości (np.: fg i fd, fg>=fd), dla których moc sygnału spada o połowę.

1. **Co to jest transmitancja filtru?**

Transmitancja filtru (charakterystyka częstotliwościowa) jest transformatą Fouriera odpowiedzi impulsowej filtru:

https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4241.gif

Moduł transmitancji https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4242.gifjest charakterystyką amplitudową filtru, zaś https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4243.gif jest charakterystyką fazową.

1. **Transformacja Fouriera**

Transformata Fouriera http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4264.gif sygnału ciągłego f(t) wyraża się następującym wzorem:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4265.gif

gdzie t - czas ciągły.

Transformacja ta przekształca dziedzinę czasu w dziedzinę widma.

Możliwe jest przekształcenie odwrotne, tj. przejście z dziedziny widma w dziedzinę czasu poprzez odwrotną transformację Fouriera:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4266.gif

W nagraniach cyfrowych dziedzina czasu zostaje poddana dyskretyzacji i zamiast ciągłej funkcji f(t) otrzymuje się sygnał {x(nT)}, gdzie T - okres próbkowania. Do analizy dźwięku stosowana jest wówczas dyskretna transformacja Fouriera.

Dyskretna transformata Fouriera X(k) dla okna czasowego o długości N definiowana jest na ciągu próbek x(0), , x((N-1)T) w sposób następujący:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4268.gif

gdzie http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4269.gif

Odwrotna dyskretna transformacja Fouriera opisywana jest poniższym wzorem:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4270.gif

Dla ciągu próbek o długości 2nopracowano szybki algorytm wyznaczania transformaty Fouriera (Fast Fourier Transform), tzw. szybką transformatę Fouriera.

Aby skorzystać z tego algorytmu, stosowane jest uzupełnianie ciągu próbek do najbliższej potęgi dwójki (zeropadding).

Własności transformacji Fouriera

Efektem ubocznym próbkowania sygnału są repliki widma, gdyż widmem ciągu impulsów próbkujących jest ciąg impulsów, a operacji mnożenia (sygnału przez ciąg impulsów próbkujących) w dziedzinie czasu odpowiada splot transformaty Fouriera sygnału i transformaty funkcji okna w dziedzinie widma:

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4271.gif

1. **Na czym polega okienkowanie sygnału i w jakim celu jest stosowane? Wymienić najpopularniejsze funkcje okienkowe**

Dla różnych długości analizowanej ramki otrzymuje się różne wyniki analiz.

Wybranie fragmentu danych o długości N oznacza, że sygnał na tym odcinku został przemnożony przez 1, zaś na pozostałych przez 0. Jest to równoważne przemnożeniu sygnału przez sygnał prostokątny o szerokości N i wysokości 1. Operację tę nazywamy okienkowaniem sygnału (z użyciem okna prostokątnego).

Operację okienkowania można zapisać jako:

v(n) = w(n) \* s(n)

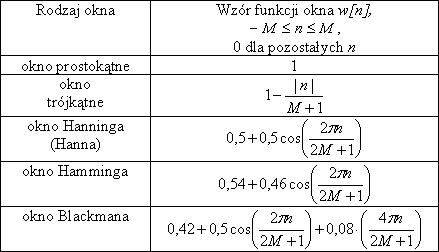
gdzie:

s(n) - sygnał wejściowy,   
v(n) - sygnał wynikowy otrzymany poprzez okienkowanie,   
w(n) - funkcja okna.

Skutkiem ubocznym okienkowania z użyciem okna prostokątnego są przecieki widma (listki boczne).

Poprzez zastosowanie okna o wartościach bliskich 0 na brzegach przedziału [0, N] możemy zmniejszyć wysokość listków bocznych - kosztem poszerzenia listka głównego i rozmycia prążków widma (tj. pogorszenia rozdzielczości).

Podstawowe funkcje okienkowe przedstawiono w poniższej tabeli.



Analiza dźwięku jest zwykle wykonywana z zastosowaniem funkcji okna różnych od prostokątnego.

1. **Różnice między transformatą Fouriera i Falkową**

Podstawową metodą analizy widmowej jest transformata Fouriera. Przebieg zmian składu częstotliwościowego dźwięku w czasie można prześledzić wykonując serię transformat Fouriera i umieszczając wyniki na wspólnym wykresie.

Metodą pozwalającą na jednoczesną analizę czasowo-częstotliwościową jest analiza falkowa.

Do analizy czasowo-częstotliwościowej stosowana jest m.in. transformata falkowa (ang. wavelet - falka).

Jest ona przekształceniem liniowym, w którym dwuwymiarowa reprezentacja sygnału za pomocą odpowiednich funkcji elementarnych pozwala na rekonstrukcję sygnału w postaci kombinacji liniowej tych funkcji.

Analiza falkowa umożliwia analizę dźwięku ze zmienną rozdzielczością (MRA - ang. multiresolution analysis). Dla dowolnej funkcjihttps://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4278.gif analiza MRA oparta jest na rozkładzie przestrzeni https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4279.gif na sumę podprzestrzeni:

https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4280.gif

gdzie:

j - poziom rozdzielczości,

https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4281.gif - przestrzenie aproksymacji (ogółów),

https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4282.gif - przestrzenie szczegółów.

Funkcja https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4283.gif rozkładana jest za pomocą funkcji https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4284.gif i https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4285.gif .

Dla najpopularniejszego typu tej analizy, tj. przy połowieniu pasma częstotliwości, oraz przy wykorzystaniu baz ortonormalnych, https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4286.gif i https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4287.gifmuszą spełniać następujące założenia:

https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4288.gif ,

https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4289.gif jest bazą ortonormalną dla Vj,

https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4290.gif ,

https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4291.gif jest bazą ortonormalną dla Wj,

g, h - współczynniki odpowiadające filtrom górno- i dolnoprzepustowemu, https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4292.gif

Funkcja https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4293.gif nazywana jest funkcją skalującą, zaś https://edux.pjwstk.edu.pl/mat/1929/lec/w5/Image4294.gif - falką macierzystą

Funkcje skalujące i macierzyste można definiować na różne sposoby w zależności od zastosowanych filtrów.

Analiza czasowo-częstotliwościowa oparta na transformacji falkowej jest obecnie coraz częściej stosowana, ze względu na ograniczoną rozdzielczość analizy FFT w dziedzinie czasu bądź częstotliwości.

1. **Wymienić najpopularniejsze falki stosowane w analizie Falkowej**

Najpopularniejsze są falki Haara, Daubechies, Meyera, Shannona, Morleta, oraz "kapelusz meksykański".

1. **Podaj definicję ciśnienia SPL.**

SPL(Sound Pressure Level) jest skalą decybelową ze ściśle określonym poziomem odniesienia.

Sound Pressure Level (SPL)=20 log (p/p0)=10 log (p/p0)2

gdzie p0=2\*10-5N/m2.

SPL jest proporcjonalny do średniego pierwiastka z amplitudy dźwięku.

1. **Dwa 80 dB źródła dodając się generują wartość mocy akustycznej równą ...?**

Łączenie 2 i więcej źródeł nie powoduje sumowania wartości dB SPL czy PWL.

2 jednakowe źródła dają zwiększenie PWL o 3dB.

2 jednakowe źródła dają 3 dB zwiększenie SPL, przy założeniu braku interferencji:

10 log 2I1/I2=10 log 2 + 10 log I1/I2

Zatem dwa 80-decybelowe źródła dodając się dają 83 dB SPL

1. **Co to są izofony?**

SPL tonu odniesienia (1 kHz) identyfikuje izofony, czyli krzywe jednakowej głośności w fonach, wyznaczanych w dB SPL. Izofony dla różnych poziomów głośności przedstawiono poniżej.

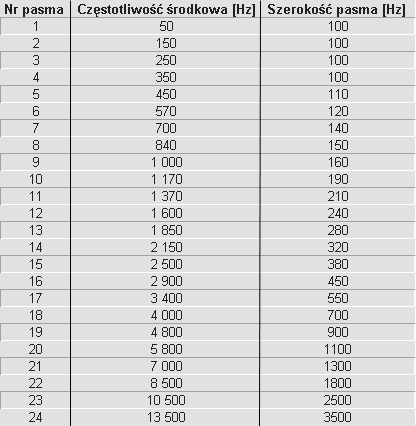
1. **Co to są pasma krytyczne?**

Pasmo krytycznejest to elementarne pasmo częstotliwości o szerokości https://edux.pjwstk.edu.pl/mat/1929/lec/w3/Image4203.gif wydzielone z ciągłego pasma szumów i zawierające w sobie moc akustyczną równą mocy akustycznej tonu prostego o częstotliwości f położonej w środku tego pasma, przy czym rozpatrywany ton prosty ma taką intensywność, że zagłuszany przez nieograniczone widmo szumów ciągłych znajduje się dokładnie na granicy słyszalności.

Jeśli ton prosty o częstotliwości f jest słyszany na tle równomiernego i nieograniczonego szumu, to efekt zagłuszania tonu przez szum wywołany jest jedynie działaniem szumów mieszczących się w paśmie krytycznym.

Pasma krytyczne są regionami "wyróżnionymi" na błonie podstawnej, wyznaczonymi przez eksperymenty psychoakustyczne. Wyróżniono około 24 pasm na błonie podstawnej. Każde pasmo ma około 1.3 mm długości i obejmuje około 1300 neuronów.

Koncepcja pasm krytycznych została wprowadzona teoretycznie przez Fletchera (12 pasm), a następnie rozwinięta eksperymentalnie przez Zwickera. Pasma krytyczne wg Zwickera (24 pasma) przedstawiono w poniższej tabeli. Podział na 24 pasma krytyczne wg Zwickera wyznacza skalę barków, gdzie 1 bark odpowiada 1 pasmu krytycznemu.



Głośność percypowana dla dźwięków złożonych jest związana z pasmami krytycznymi.

Gdy szerokość pasma dźwięku złożonego jest mniejsza od szerokości pasma krytycznego, wtedy głośność jest w przybliżeniu niezależna od szerokości pasma. Dźwięk taki oceniany jest jako jednakowo głośny z tonem o takim samym natężeniu i o częstotliwości równej częstotliwości środkowej pasma tego dźwięku. Jeśli szerokość pasma dźwięku będzie zwiększana ponad szerokość pasma krytycznego, głośność tego dźwięku zacznie wzrastać.

Dla dźwięków o szerszym paśmie głośność rośnie wraz z pasmem i następuje tzw. sumowanie głośności. Np. szerokopasmowy dźwięk orkiestry grającej akord wydaje się być głośniejszy niż pojedynczy dźwięk fletu, nawet, jeśli dźwięki mają ten sam poziom SPL.

1. **Podaj prawo Webera-Fechnera**

Prawo Webera-Fechnera mówi, iż percepcja w postępie arytmetycznym zależy od geometrycznego postępu faktów fizycznych.

Przykładowo, percepcja równomiernego przyrostu stopni szarości w przejściu od czerni do bieli wymaga geometrycznego przyrostu luminancji.

1. **Wymienić kilka częstotliwości próbkowania stosowanych w cyfrowych systemach audio**

Częstotliwości próbkowania stosowane w cyfrowych systemach audio są następujące:

* 5500 Hz - standard Macintosh (=44100/8)
* 7333 Hz (=1/6 częstotliwości Maca)
* 8000 Hz - standard telefoniczny do kodowania μ-law, a-law
* 8012.8210513 Hz - standard NeXT, używany z kodekiem Telco
* 11025 Hz (=22050/2)
* 16000 Hz - standard telefoniczny G.722
* 16726.8 Hz - standard telewizyjny NTSC (= 7159090.5/(214\*2))
* 18900 Hz - standard CD-ROM
* 22050 Hz - standard Macintosha, CD/2
* 22254.[54] - standard złącza monitora MacIntosh 128k
* 32000 Hz - standard DAB (Digital Audio Broadcasting), NICAM (Nearly-Instantaneous Companded Audio Multiplex) - np. BBC; inne systemy TV, HDTV, R-DAT
* 32768 Hz (32\*1024)
* 37800 Hz - high quality CD-ROM
* 44056 Hz - częstotliwość próbkowania używana w sprzęcie profesjonalnym (kompatybilność z NTSC)
* 44100 Hz - CD audio - najpopularniejsza częstotliwość w aplikacjach profesjonalnych i domowych
* 48000 Hz - R-DAT
* 49152 Hz (48\*1024)
* powyżej 50000 Hz - używane niekiedy w profesjonalnych systemach cyfrowego przetwarzania sygnałów
* 96000 Hz - high resolution R-DAT

1. **Na czym polega procedura przepróbkowania i w jakim celu jest stosowana?**

W praktyce wybór formatu cyfrowego zapisu dźwięku jest zwykle kompromisem między jakością sygnału a oszczędnością pamięci. Pliki przeznaczone do przechowywania szerokopasmowego sygnału muzycznego wysokiej jakości będą narzucać większe wymagania sprzętowe/softwareowe, zaś pliki do przechowywania sygnału mowy z ograniczeniem pasma częstotliwości i dynamiki będą mogły być rejestrowane przy słabszych parametrach.

Wraz z rozwojem techniki cyfrowego zapisu sygnałów akustycznych i stosowaniem systemów komputerowych do obróbki dźwięku, powstało wiele różnych formatów składowania danych - plików dźwiękowych. Ich znaczna ilość spowodowana jest głównie różnorodnością systemów operacyjnych.

Ponieważ metody akwizycji i przechowywania danych dźwiękowych zostały opracowane niezależnie przez wielu producentów sprzętu i oprogramowania, stąd powstało wiele różnych formatów plików dźwiękowych. Z powyższym wiąże się zagadnienie konwersji formatów plików dźwiękowych.

Przy obsłudze komputerowych systemów obróbki dźwięku niezbędna jest znajomość podstawowych formatów plików dźwiękowych, ich dostępności systemowej i przeznaczenia.

Znajomość tej problematyki jest szczególnie istotna przy przenoszeniu plików dźwiękowych pomiędzy systemami operacyjnymi.

Przenoszenie plików dźwiękowych pomiędzy systemami umożliwiają programowe konwertery formatów. Programy te umożliwiają niekiedy także proste przekształcenia sygnału np. dodanie echa, odwrócenie przebiegu, decymację, zmianę amplitudy itp.

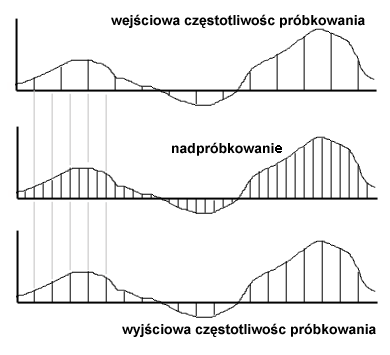
Konwersja formatów plików dźwiękowych wymaga uwzględnienia 2 zagadnień:

* Konwersja parametrów zapisu: częstotliwości próbkowania, liczby kanałów, rozdzielczości bitowej,
* Konwersja typów plików (formatu zapisu danych w pliku).

Konwersja częstotliwości próbkowania jest procedurą dwuetapową, przebiegającą w następującej kolejnośći:

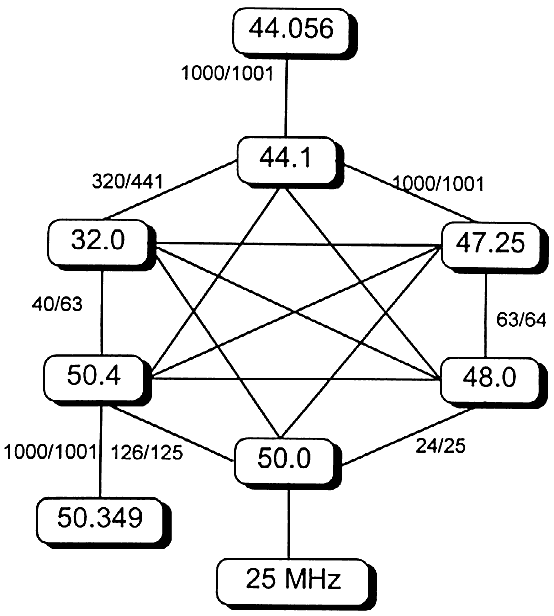
* Nadpróbkowanie (ang. oversampling) - generowanie dodatkowych próbek
* Usuwanie nadmiarowych próbek

Częstotliwość nadpróbkowania powinna być NWW (najmniejszą wspólną wielokrotnością) źródłowej i docelowej częstotliwości próbkowania. Procedurę tę ilustruje poniższy rysunek.



(źródło: K. Szklanny, Multimedia, materiały do wykładów, 2004 -[kszklanny@pjwstk.edu.pl](mailto:-kszklanny@pjwstk.edu.pl))

Relacje między najpopularniejszymi częstotliwościami próbkowania przedstawia poniższy schemat.



(źródło: Czyżewski A., Dźwięk cyfrowy, Akademicka Oficyna Wydawnicza EXIT, Warszawa 1998)

Procedura nadpróbkowania wymaga wygenerowania dodatkowych próbek.

Dodatkowe próbki generowane są za pomocą różnego typu algorytmów interpolacyjnych.

1. **Wymienić najpopularniejsze sposoby kodowania dźwięku**

Sposoby kodowania (zapisu) fali dźwiękowej są następujące:

* PCM
* DPCM
* ADPCM
* Kodeki kompandorowe:
  + Mu-law (standard amerykańsko-japoński)
  + law (standard europejski)
* Kodeki źródła - wokodery
* Kodeki hybrydowe

gdzie kodek = koder + dekoder.

1. **Co to jest wokoder i do czego jest stosowany?**

Wokoder(Voice Coder) jest kodekiem źródła, przewidziany do transmisji sygnału mowy.

Kodek źródła tworzy model źródła dźwięku i dokonuje rekonstrukcji sygnału na podstawie tego modelu.

Używane są 2 podstawowe modele sygnału:

* Dźwięczny (pobudzenie tonowe),
* Bezdźwięczny (pobudzenie szumowe).

Zaletą wokodera jest to, że sygnał przekazywany jest w bardzo małym pliku.

Wadą wokodera jest, iż nadaje się on do kodowania jedynie określonego typu sygnałów - nie nadaje się do kodowania np. muzyki.

1. **Wymienić najważniejsze formaty plików dźwiękowych**

Główne formaty plików dźwiękowych to:

* .snd, .au (NeXT, Sun)
* .wav (Microsoft, IBM)
* .mp3
* .mid (MIDI)

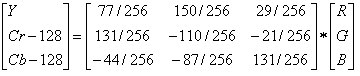
a także

* aiff
* mp3pro
* mp4
* ogg
* vqf
* qt (QuickTime audio)
* ra, rm, ram (RealAudio)
* wma
* rmf
* Liquid Audio
* a2bmusic

Niektóre formaty/standardy oparte są na kompresji perceptualnej, kodującej stratnie dźwięk w sposób dostosowany do charakterystyki słuchu człowieka: MPEG, AC-3 (HDTV), PASC (magnetofon DCC).

1. **Wymienić najważniejsze sposoby kodowania zastosowane w kompresji wg standardu JPEG**

Przekształcenie obrazu RGB w YCrCb wg poniższego wzoru



Kolory RGB skwantowane na 220 poziomach zostają zamienione na luminancję (jaskrawość) Y i chrominancję (kolorowość) CrCb, również 220 poziomów.

2. Kodowana jest 1 para wartości chrominancji na każde 2 wartości luminancji.

3. Zastosowanie DCT (Discrete Cosine Transform) dla bloków 8x8 pikseli

4. Kwantyzacja, zależna od częstotliwości przestrzennej.

5. Kodowanie RLE (Run Length Encoding) i metodą Huffmana, w oparciu o obliczanie entropii i przewidywanie oczekiwanego wzorca danych.

JPEG wykorzystuje względną niewrażliwość ludzkiego oka na kontrasty koloru (odcienie), tj. zmiany chrominancji, w porównaniu z luminancją. Możliwa jest więc zmiana kroku kwantyzacji dla każdego składnika częstotliwości, tj. większy krok może reprezentować mniej znaczące częstotliwości

1. **Jakie są główne zalety i wady kompresji JPEG? Co jest ich przyczyną?**

Poniżej przedstawiono obrazy spakowane kompresją JPEG, także w porównaniu z oryginałem, aby pokazać wady, jakie mogą się pojawić na obrazie w wyniku kompresji JPEG.

Zaletą kompresji JPEG jest duży stopień upakowania plików. Kompresja JPEG ma też swoje wady, widoczne zwłaszcza przy większej kompresji:

* Efekt zblokowania pikseli (ang. blocking),
* Efekt zniekształcenia krawędzi (ang. fringing).

Efekt zblokowania pikseli może się uwidocznić szczególnie na dużych powierzchniach w jednym kolorze z płynną zmianą odcieni, np. na zdjęciach nieba. Efekt ten zilustrowany jest również na poniższych zdjęciach.

Efekt zblokowania pikseli jest widoczny na płatkach kwiatu.

Poniżej przedstawiono przykład zdjęć zakodowanych w formacie JPEG, gdzie widoczne jest rozmycie krawędzi.

Na dolnym zdjęciu zniekształcenia krawędzi widoczne są szczególnie na górnej krawędzi telewizora.

1. **Jakie techniki kompresji zastosowano w standardzie MPEG?**

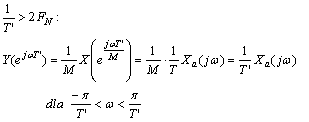
Techniki kompresji zastosowane w MPEG to:

* Discrete Cosine Transform (DCT),
* Kwantyzacja,
* Kodowanie Huffmana,
* Kodowanie predykcyjne - obliczanie różnic między ramkami, a następnie kodowanie wyłącznie tych różnic,
* Predykcja dwustronna - na podstawie obrazów poprzednich i następnych.

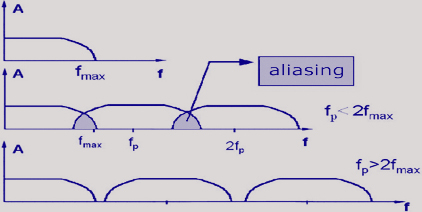
1. **Warunek Nyquista, częstotliwość Nyquista**

Aby uniknąć aliasingu, czyli nakładania replik widma, należy usunąć z sygnału częstotliwości powyżej połowy częstotliwości próbkowania (częstotliwości Nyquista) http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4273.gif, gdzie http://edu.pjwstk.edu.pl/wyklady/mul/scb/w5/Image4274.gif- częstotliwość próbkowania.

Aby uniknąć aliasingu (nakładania widma), nadpróbkowany sygnał nie może zawierać częstotliwości powyżej częstotliwości Nyquista, czyli powyżej połowy docelowej częstotliwości próbkowania:



Przy cyfrowej rejestracji dźwięku należy zadbać o to, aby w rejestrowanym sygnale nie było częstotliwości powyżej połowy częstotliwości próbkowania, http://edu.pjwstk.edu.pl/wyklady/mul/scb/w12/Image4492.gif. W przeciwnym razie repliki widma dźwięku będą się nakładać, co nazywamy aliasingiem (patrz rysunek poniżej).



Warunkiem koniecznym zapobieżenia aliasingowi jest

http://edu.pjwstk.edu.pl/wyklady/mul/scb/w12/Image4494.gif

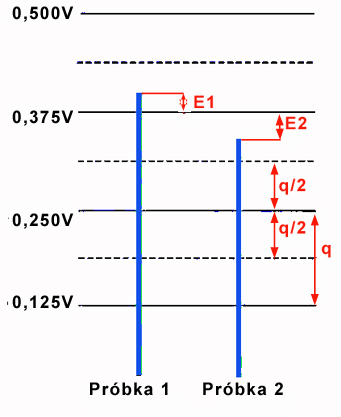
gdzie http://edu.pjwstk.edu.pl/wyklady/mul/scb/w12/Image4495.gif- częstotliwość próbkowania, http://edu.pjwstk.edu.pl/wyklady/mul/scb/w12/Image4496.gif- maksymalna częstotliwość występująca w sygnale.

Przy ustalonej częstotliwości próbkowania niezbędne jest ograniczenie pasma sygnału wejściowego poprzez filtrację antyaliasingową dolnoprzepustową.

Aby uniknąć aliasingu stosuje się zatem przed rejestracją filtrację dolnoprzepustową z częstotliwością odcięcia http://edu.pjwstk.edu.pl/wyklady/mul/scb/w12/Image4497.gif(częstotliwość Nyquista).

1. **Na czym polega kwantowanie? Kwantyzacja wielobitowa, liniowa, nieliniowa**

Kwantyzacja (wielobitowa) polega na podziale osi amplitudy na pewną liczę poziomów (zależnie od rozdzielczości bitowej), a następnie przypisanie wartości amplitudy w każdej próbce jak na poniższym rysunku.



(źródło: K. Szklanny, Multimedia, materiały do wykładów, 2004 -[kszklanny@pjwstk.edu.pl](mailto:-kszklanny@pjwstk.edu.pl))

Obie próbki na powyższym rysunku zostały skwantowane na poziomie 0.375V z błędami https://edux.pjwstk.edu.pl/mat/1929/lec/w12/Image4502.gifi https://edux.pjwstk.edu.pl/mat/1929/lec/w12/Image4503.gif.

Im większa liczba przedziałów kwantyzacji w stosunku do amplitudy sygnału, tym mniejsze błędy kwantyzacji.

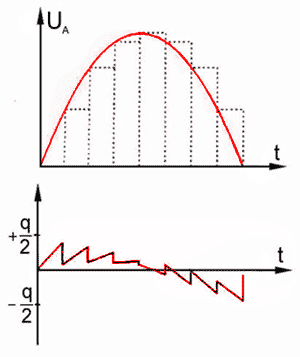
Liczba poziomów kwantyzacji wynosi https://edux.pjwstk.edu.pl/mat/1929/lec/w12/Image4504.gif, gdzie n - długość słowa zapisu.

Długości słów zostały znormalizowane i wynoszą 8, 14, 16, 20 i 24 bity.

Szerokość przedziału kwantyzacji (rozdzielczość) wynosi https://edux.pjwstk.edu.pl/mat/1929/lec/w12/Image4505.gif.

Rozdzielczość bitowa wyznacza dynamikę, która wynosi https://edux.pjwstk.edu.pl/mat/1929/lec/w12/Image4506.gif[dB]

Wskutek nakładania błędów kwantyzacji na sygnał przy przetwarzaniu C/A powstaje szum kwantyzacji, zilustrowany na poniższym rysunku.



(źródło: K. Szklanny, Multimedia, materiały do wykładów, 2004 -[kszklanny@pjwstk.edu.pl](mailto:-kszklanny@pjwstk.edu.pl))

Szum kwantyzacji jest szumem białym.

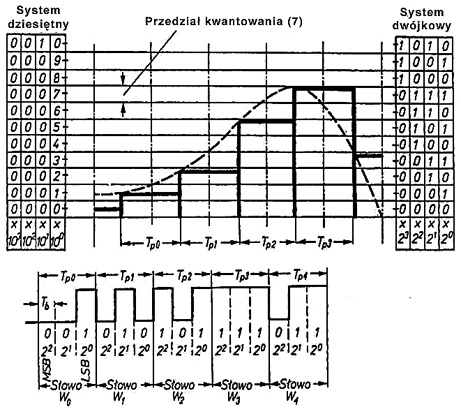
Moc szumu kwantyzacji wynosi https://edux.pjwstk.edu.pl/mat/1929/lec/w12/Image4508.gif.

Odstęp szumu kwantyzacji od sygnału wynosi 20log Umax/Uq = 6n + 1,8 [dB], gdzie Uq - wartość skuteczna szumu kwantyzacji (RMS - Root Mean Square), Umax - napięcie maksymalne.

W ujęciu matematycznym wartość skuteczna prądu zmiennego to jego wartość szczytowa podzielona przez https://edux.pjwstk.edu.pl/mat/1929/lec/w12/Image4509.gif. RMS w odniesieniu do wielkości elektrycznych odpowiada wartości skutecznej np. prądu lub mocy.

Gdy poziom sygnałów kwantowanych przewyższa poziom najwyższego poziomu kwantowania, powstają zniekształcenia nieliniowe, zilustrowane na poniższym rysunku. Rozwiązaniem tego problemu jest pozostawienie podczas nagrywania rezerwy około 6 dB na wysterowanie.

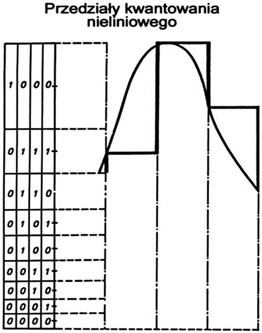
Omówiona dotychczas kwantyzacja jest kwantyzacją liniową (PCM - pulse-code modulation), co ilustruje poniższy rysunek.



(źródło: K. Szklanny, Multimedia, materiały do wykładów, 2004 -[kszklanny@pjwstk.edu.pl](mailto:-kszklanny@pjwstk.edu.pl))

W kwantyzacji liniowej następuje dzielenie zakresu zmian sygnału analogowego na równe przedziały.

Kwantyzacja nieliniowa polega na podziale zakresu zmian sygnału analogowego w nierównych odstępach, co ilustruje poniższy rysunek.



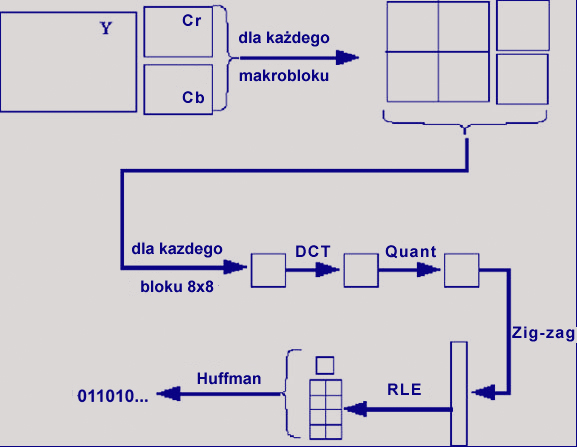
1. **Ile poziomów kwantowania daje zastosowanie słowa 8-bitowego?**

Im większa liczba przedziałów kwantyzacji w stosunku do amplitudy sygnału, tym mniejsze błędy kwantyzacji.

Liczba poziomów kwantyzacji wynosi https://edux.pjwstk.edu.pl/mat/1929/lec/w12/Image4504.gif, gdzie n - długość słowa zapisu.

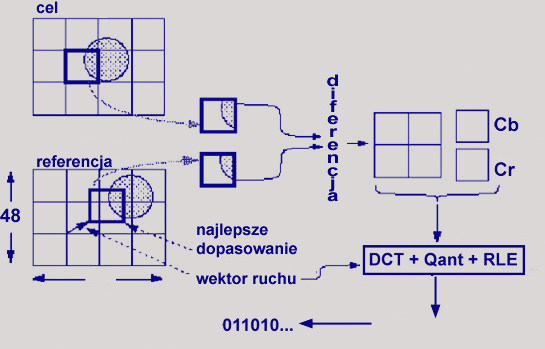
1. **Proszę przedstawić schemat kodowania makrobloku typu I.**

Schemat kodowania ramki typu I przedstawia poniższy rysunek.



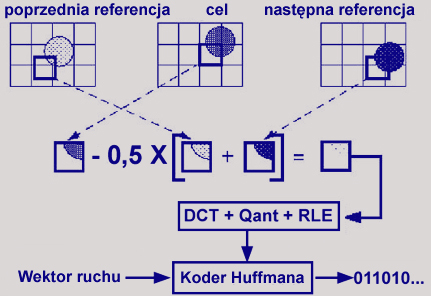
1. **Proszę przedstawić schemat kodowania makrobloku typu P.**

Schemat kodowania ramki typu P przedstawiono poniżej.



1. **Proszę przedstawić schemat kodowania makrobloku typu B.**

Schemat kodowania ramki typu B przedstawia poniższy rysunek.



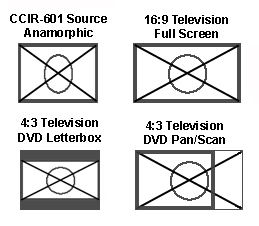
1. **Jaki jest zakres pojemności DVD? Rodzaje płyt DVD i ich pojemności**

W zależności od liczby stron i warstw płyt DVD wyróżniamy ich następujące rodzaje:

* DVD-5: jednnostronna, jednowarstwowa (single-sided, single layer), 4,7 GB
* DVD-10: dwustronna, jednowarstwowa (double-sided, single layer), 9,4 GB
* DVD-9: jednnostronna dwuwarstwowa (single-sided with dual layers), 8,5 GB
* DVD-18: dwustronna dwuwarstwowa (double-sided with dual layers), 17 GB.

1. **Proszę przedstawić sposoby dopasowania obrazów w proporcji 16:9 do wyświetlania w proporcje 4:3.**

Proporcje obrazu w DVD można wybrać w zilustrowanych poniżej wersjach.



We wczesnych filmach stosowano proporcje 4:3, podobnie na taśmie 35mm czy w telewizji, zarówno PAL jak i NTSC. Z rozwojem filmu pojawiły się szersze proporcje, co wymaga dostosowywania proporcji przy przenoszeniu filmu na ekran komputerowy lub klasycznego telewizora. Standardem dla telewizji wysokiej rozdzielczości HDTV i telewizorów szerokoekranowych jest 16:9. Najbardziej rozpowszechnione proporcje kinowe to 1.85:1 i 2.35:1.

1. **Proszę wymienić formaty zapisu audio stosowane na dyskach DVD**

Dźwięk w DVD jest wysokiej jakości, wielokanałowy:

* Do 8 kanałów audio 48kHz,
* Do 4 kanałów audio 96kHz.

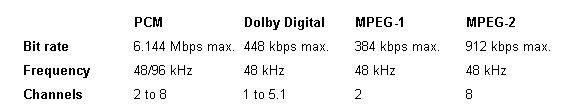
Formaty stosowane w DVD Audio to:

* PCM (pulse code modulated),
* Dolby Digital,
* MPEG.

W krajach NTSC stosowane są formaty PCM i Dolby Digital, zaś MPEG jako opcja.

W krajach stosujących PAL używane są formaty PCM i MPEG, zaś Dolby Digital jako opcja.

Formaty audio w DVD ilustruje poniższy rysunek.



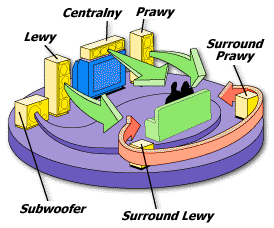
1. **System Dolby Digital (AC-3)**

Dolby Digital/Dolby Surround AC-3

AC-3 jest cyfrowym systemem zapisu dźwięku 5.1, z następującymi kanałami:

* Lewy kanał przedni,
* Centralny kanał przedni,
* Prawy kanał przedni,
* Lewy kanał Surround,
* Prawy kanał Surround,
* Suwboofer.

Rozmieszczenie głośników ilustruje poniższy rysunek.



System ten umożliwia odgrywanie dźwięku poprzez niższe systemu w przypadku, gdy użytkownik nie posiada dekodera Dolby Digital. Dźwięk może być odgrywany w trybach mono, stereo, Dolby Pro-Logic, oraz w pełnym sześciokanałowym Dolby Digital.

1. **Jakie metadane można zastosować w opisie baz audio?**

Metadane do reprezentacji zawartości audio to np. wykonawcy, nuty, czy tekst.

Algorytmiczna ekstrakcja cech dźwięku może wydobyć takie atrybuty jak natężenie dźwięku, głośność, wysokość dźwięku, jasność dźwięku, czy parametry statystyczne, np. korelacja, wariancja etc..

1. **Na czym polega konstrukcja drzew czwórkowych typu MX-quastrees i gdzie są one stosowane?**

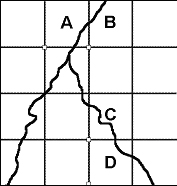
W przypadku drzew czwórkowych typu MX zakłada się, że na mapę naniesiona jest siatka 2kx 2k.

Kształt drzewa (i jego wysokość) jest niezależna od liczby węzłów w drzewie.

Punktem podziału w danym węźle jest zawsze środek obszaru reprezentowanego przez ten węzeł.

Wszystkie punkty reprezentowane są na poziomie liści, co upraszcza ewentualne usuwanie ich z drzewa.

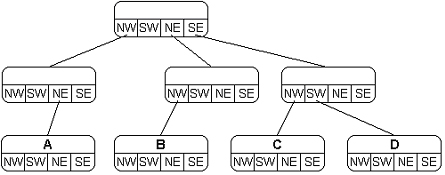
Podział mapy dla drzewa MX w węźle N przedstawiono poniżej.



(źródło: V.S. Subrahmanian: Principles of Multimedia Database Systems, Morgan Kaufmann 1998)

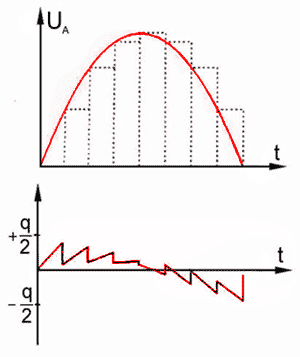
Załóżmy, ze do drzewa MX należy wstawić obszary A, B, C i D.

Otrzymane drzewo dla tego przykładu przedstawiono na poniższym rysunku.



1. **W wyniku czego powstaje szum kwantyzacji?**

Wskutek nakładania błędów kwantyzacji na sygnał przy przetwarzaniu C/A powstaje szum kwantyzacji, zilustrowany na poniższym rysunku.



(źródło: K. Szklanny, Multimedia, materiały do wykładów, 2004 -[kszklanny@pjwstk.edu.pl](mailto:-kszklanny@pjwstk.edu.pl))

Szum kwantyzacji jest szumem białym.

Moc szumu kwantyzacji wynosi http://edu.pjwstk.edu.pl/wyklady/mul/scb/w12/Image4508.gif.

Odstęp szumu kwantyzacji od sygnału wynosi 20log Umax/Uq = 6n + 1,8 [dB], gdzie Uq - wartość skuteczna szumu kwantyzacji (RMS - Root Mean Square), Umax - napięcie maksymalne.

W ujęciu matematycznym wartość skuteczna prądu zmiennego to jego wartość szczytowa podzielona przez http://edu.pjwstk.edu.pl/wyklady/mul/scb/w12/Image4509.gif. RMS w odniesieniu do wielkości elektrycznych odpowiada wartości skutecznej np. prądu lub mocy.

Gdy poziom sygnałów kwantowanych przewyższa poziom najwyższego poziomu kwantowania, powstają zniekształcenia nieliniowe, zilustrowane na poniższym rysunku. Rozwiązaniem tego problemu jest pozostawienie podczas nagrywania rezerwy około 6 dB na wysterowanie.

1. **Na czym polega zjawisku maskowania dźwięku?**

Jednym z perceptualnych aspektów słyszenia jest zjawisko maskowania.

Głośniejszy dźwięk może sprawić, że inny, cichszy dźwięk nie będzie słyszalny (głośniejszy zamaskuje cichszy), zależnie od wzajemnych relacji częstotliwości i głośności obu dźwięków.

Czyste tony o bliskich częstotliwościach maskują się bardziej niż oddalone.

Czysty ton o wyższej częstotliwości jest maskowany efektywniej niż o niższej.

Im większa intensywność tonu maskującego (maskera), tym szersze pasmo maskowania.

Najistotniejszym faktem psychoakustycznym jest efekt maskowania.

Maskowanie polega na zagłuszaniu dźwięku/ów przez inne dźwięki. Maskowanie równoczesne polega na tym, że w zależności od wzajemnego natężenia i częstotliwości pewne tony stają się niesłyszalne (zostają zagłuszone, zamaskowane) w obecności innych tonów w ich sąsiedztwie, co ilustruje poniższy rysunek. Dźwięk maskujący nazywany jest maskerem.

Oprócz maskowania równoczesnego zachodzi także maskowanie nierównoczesne. Premaskowanie (maskowanie wstecz, do kilkunastu ms)

związane jest z szybszym przetwarzaniem przez układ słuchowy dźwięków głośniejszych, zatem dźwięk późniejszy, ale głośniejszy może zamaskować dźwięk wcześniejszy, ale cichszy. Postmaskowanie (maskowanie wprzód, do kilkuset ms) związane jest z relaksacją neuronów i zależy od głośności oraz czasu trwania tonu maskującego. Zatem dźwięk może zostać zamaskowany przez dźwięk wcześniejsze, w znacznie większym zakresie niż dla efektu premaskowania.

Dla każdego tonu w sygnale audio można wyznaczyć próg maskowania. Jeśli jakiś ton leży poniżej tego progu, zostanie zamaskowany przez ton głośniejszy i nie będzie słyszalny. Niesłyszalne elementy sygnału audio mogą zostać wyeliminowane przez koder.