

informatyka+

Algorytmika i programowanie

Bazy danych

Multimedia, grafika i technologie internetowe

Sieci komputerowe

Tendencje w rozwoju informatyki i jej zastosowań

informatyka+

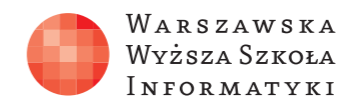
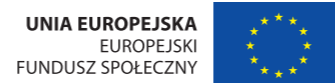
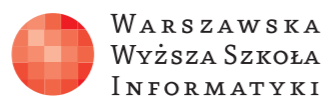
Wszechnica Informatyczna: Multimedia, grafika i technologie internetowe

Treści multimedialne
– kodowanie,
przetwarzanie, prezentacja

Andrzej Majkowski

Człowiek – najlepsza inwestycja

Człowiek – najlepsza inwestycja



Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

Treści multimedialne – kodowanie, przetwarzanie, prezentacja



Rodzaj zajęć: Wszelchnica Informatyczna

Tytuł: Treści multimedialne – kodowanie, przetwarzanie, prezentacja

Autor: dr inż. Andrzej Majkowski

Redaktor merytoryczny: prof. dr hab. Maciej M Sysło

Zeszyt dydaktyczny opracowany w ramach projektu edukacyjnego **Informatyka+** – ponadregionalny program rozwijania kompetencji uczniów szkół ponadgimnazjalnych w zakresie technologii informacyjno-komunikacyjnych (ICT).

www.informatykaplus.edu.pl

kontakt@informatykaplus.edu.pl

Wydawca: Warszawska Wyższa Szkoła Informatyki

ul. Lewartowskiego 17, 00-169 Warszawa

www.wysi.edu.pl

rektorat@wysi.edu.pl

Projekt graficzny: FRYCZ I WICHA

Warszawa 2010

Copyright © Warszawska Wyższa Szkoła Informatyki 2010

Publikacja nie jest przeznaczona do sprzedaży.



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



WARSZAWSKA
WYŻSZA SZKOŁA
INFORMATYKI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

Treści multimedialne – kodowanie, przetwarzanie, prezentacja



Andrzej Majkowski

Politechnika Warszawska

amajk@ee.pw.edu.pl

Streszczenie

Przedmiotem tych zajęć są następujące zagadnienia:

1. Projekt prezentacji multimedialnej z użyciem tekstu, obrazu, dźwięku. Przedstawienie możliwości programów Microsoft Power Point, Open Office Impress. Tworzenie prezentacji w formacie PDF; program Adobe Reader.
2. Rejestracja sygnałów dźwiękowych przy użyciu różnych typów mikrofonów. Systemy odsłuchu dźwięku. Standardy zapisy dźwięku. Jakość sygnału dźwiękowego. Kodowanie MP3. Podstawowe efekty dźwiękowe.
3. Rejestrowanie sygnałów wizyjnych, kamery. Podstawy tworzenia filmów, authoring.
4. Wyświetlacze obrazów – rodzaje, ich wady i zalety (CRT, LCD, projektory). Ocena jakości obrazu. Konfiguracje sprzętowe zestawów multimedialnych, kino domowe.

Sprzęt i oprogramowanie potrzebne do przeprowadzenia kursu

- komputer klasy PC, system Windows XP, nagrywarka DVD,
- pakiet OpenOffice z programem Impress, Microsoft Office Power Point, Adobe Reader,
- oprogramowanie Audacity, Windows Movie Maker, Pinnacle Studio,
- mikrofon elektretowy, dynamiczny, pojemnościowy (po 2szt.),
- wzmacniacz mikrofonowy,
- prosty mikser kanałów,
- zestaw kina domowego,
- telewizor CRT, LCD, rzutnik multimedialny, odtwarzacz DVD.



Spis treści

| | |
|---------------------------------------------------------------------------------------|----|
| 1. Multimedia w przekazie informacji | 6 |
| 1.1 Historia terminu multimedia | 6 |
| 1.2 Dźwięk | 6 |
| 1.2.1 Zakres słyszalności | 7 |
| 1.2.2 Ocena jakości dźwięku | 8 |
| 1.3 Obraz | 8 |
| 1.3.1 Modele barw..... | 10 |
| 1.4 Zastosowanie dźwięku w multimediami | 11 |
| 1.4.1 Karty dźwiękowe | 11 |
| 1.4.2 Budowa kart dźwiękowych | 13 |
| 1.5 Prezentacja multimedialna | 14 |
| Zadanie 1. Utworzenie prezentacji multimedialnej | 15 |
| 2. Kodowanie i przetwarzanie sygnałów dźwiękowych | 18 |
| 2.1 Rejestracja sygnałów dźwiękowych..... | 18 |
| 2.1.1 Podział akustyczny mikrofonów | 18 |
| 2.1.2 Podział mikrofonów ze względu na rodzaj przetwornika (podział elektryczny)..... | 19 |
| 2.1.3 Parametry mikrofonów | 20 |
| 2.1.4 Techniki mikrofonowe..... | 21 |
| Zadanie 2. Nagrania z użyciem różnych typów mikrofonów | 21 |
| 2.2 Formaty zapisu i przechowywania plików multimedialnych | 22 |
| 2.3 Cyfrowe przetwarzanie sygnałów dźwiękowych | 24 |
| Zadanie 3. Przetwarzanie i montaż sygnału dźwiękowego | 25 |
| 2.4 Psychoakustyka – podstawowe zasady kompresji sygnałów dźwiękowych | 27 |
| 2.5 Idea kompresji MP3..... | 29 |
| 2.6 Kodowanie dźwięku w standardzie MP3 | 30 |
| 2.7 Łączenie kanałów zapisu stereofonicznego | 33 |
| 2.8 Zalety i wady standardu MP3 | 33 |
| Zadanie 4. Formaty zapisu plików muzycznych, ocena jakości nagrania | 34 |
| 3. Obrazy ruchome | 35 |
| 3.1 Komputerowa edycja obrazu | 36 |
| Zadanie 5. Pozyskiwanie materiału wideo | 37 |
| 3.2 Standardy kodowania sekwencji wideo MPEG1 i MPEG2 | 38 |
| 3.2.1 Przebieg procesu kodowania w standardach MPEG1 i MPEG2..... | 40 |
| 3.2.2 Różnice pomiędzy standardami MPEG1 i MPEG2 | 43 |
| Zadanie 6. Edycja sekwencji wizyjnej | 44 |
| 4. Odtwarzanie treści multimedialnych | 46 |
| 4.1 Telewizja analogowa i cyfrowa..... | 46 |
| 4.2 Standardy telewizji kolorowej HDTV | 46 |
| 4.3 Cyfrowa telewizja systemu DVB | 47 |
| 4.4 Poprawa jakości obrazu | 47 |
| 4.5 Kino domowe | 47 |
| 4.6 Akustyka pomieszczenia | 49 |
| Zadanie 7. Pasma przenoszenia głośników, krzywa słyszalności..... | 50 |
| 4.7 Wyświetlacz LCD | 51 |
| 4.8 Ekran plazmowy | 56 |
| 4.9 Projektory multimedialne | 59 |
| Zadanie 8. Eksport materiału wideo do formatu MPEG-2 | 61 |
| Literatura | 62 |



1. MULTIMEDIA W PRZEKAZIE INFORMACJI

Multimedia (z łac. *Multum + Medium*) to media, które wykorzystują różne formy informacji oraz różne formy ich przekazu (np. tekst, dźwięk, grafikę, animację, wideo) w celu dostarczania odbiorcom informacji lub rozrywki. Termin multimedia ma również zastosowanie w mediach elektronicznych, służących do rejestrowania oraz odtwarzania treści multimedialnych. Multimedia mają cechy tradycyjnych technik mieszanych i sztuk pięknych, jednak mają szerszy zakres. Termin bogate media jest synonimem terminu multimedia interaktywne.

1.1 HISTORIA TERMINU MULTIMEDIA

W 1965 roku termin multimedia został użyty do określenia *Exploding Plastic Inevitable*, przedstawienia Andiego Warhola, które było połączeniem muzyki rockowej, kina, eksperymentalnego oświetlenia. W ciągu następných czterdziestu lat słowo to nabierało innych znaczeń. Pod koniec lat 70. XX wieku określało prezentacje składające się z kilku jednoczesnych pokazów slajdów zgranych w czasie ze ścieżką dźwiękową. W latach 90. XX wieku nabrało swojego obecnego znaczenia. W mowie potocznej termin multimedia odnosi się do kombinacji mediów elektronicznych składających się z filmu, obrazów, dźwięków i tekstu, która umożliwia odbiór interaktywny. Wiele informacji dostępnych obecnie w Internecie odpowiada tej definicji.

Termin multimedia jest także niejednoznaczny. Treść statyczną (np. książkę papierową) można traktować jako multimedia, jeśli zawiera zarówno obrazy jak i tekst, albo jako formę interaktywną, jeśli użytkownik przewraca strony, kiedy chce. Termin wideo, w terminologii multimedialnej jest także niejednoznaczny. Jako wideo często opisuje się format pliku, format dostarczania lub format prezentacji, a nie formę treści informacji, takiej jak obrazy statyczne i ruchome. Wielorakie formy treści informacji często nie są uważane za multimedia, jeśli nie zawierają nowoczesnych form prezentacji jak audio czy wideo. Również pojedyncze formy treści informacji przetwarzające tę informację za pomocą pojedynczej metody (na przykład nieinteraktywne audio) są często nazywane multimediami aby rozróżnić media statyczne i media dynamiczne.

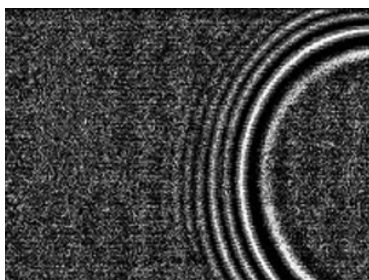
Najczęściej wykorzystywanymi w multimediami formami przekazu są, obok tradycyjnego tekstu, dźwięk i obraz (grafika, animacja, wideo). Te dwie formy będą dokładniej omówione w dalszej części opracowania.

1.2 DŹWIĘK

Fala dźwiękowa rozchodzi się jako podłużna fala akustyczna w danym ośrodku sprężystym: gazie, płynie (rys. 1). W ciałach stałych, takich jak metale, występuje również fala poprzeczna. Najczęściej mówimy o rozchodzeniu się dźwięku w powietrzu. Dźwięk, jako drgania cząsteczek, charakteryzuje się tym, że cząsteczka pobudzona przekazuje energię cząsteczce sąsiedniej, a sama drga wokół własnej osi. Skutkiem tego są lokalne zmiany ciśnienia ośrodka rozchodzące się falowo. Co ciekawe, w wodzie dźwięk rozchodzi się znacznie szybciej niż w powietrzu, a w próżni oczywiście nie rozchodzi się w ogóle. W potocznym znaczeniu **dźwięk** to każde rozpoznawalne przez człowieka pojedyncze wrażenie słuchowe.

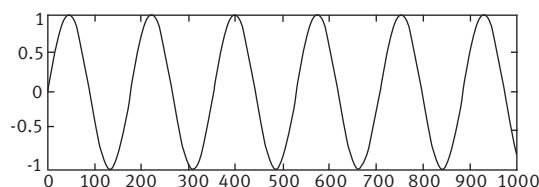
Zakres częstotliwości od 20Hz do 20kHz jest zakresem częstotliwości słyszalnych (fonicznych, audio). Dźwięki o częstotliwości mniejszej od 20Hz są nazywane **infradźwiękami**, zaś o częstotliwości większej od 20kHz – **ultradźwiękami**.

Elementarnym rodzajem dźwięku, dla którego fala dźwiękowa ma postać sinusoidy (rys. 2), jest **ton**. **Wysokość tonu** to atrybut wrażenia słuchowego umożliwiający na uszeregowanie dźwięków na skali niskie-wysokie. Przez **wysokość dźwięku** rozumie się częstotliwość drgań fali akustycznej, im wyższa częstotliwość drgań tym „wyższy” dźwięk. Dźwięki są najczęściej sygnałami złożonymi (występuje w nich wiele składowych sinusoidalnych o różnych amplitudach i częstotliwościach). Wysokość dźwięku, jest często utożsamiana z częstotliwością i zależy od niej w największym stopniu, ale nie wyłącznie. Innymi wyznacznikami wrażenia wysokości są m.in. natężenie dźwięku, czy współobecności innych tonów. Występują też różnice w postrzeganiu wysokości dźwięku między lewym i prawym uchem.



Rysunek 1.
Fala dźwiękowa

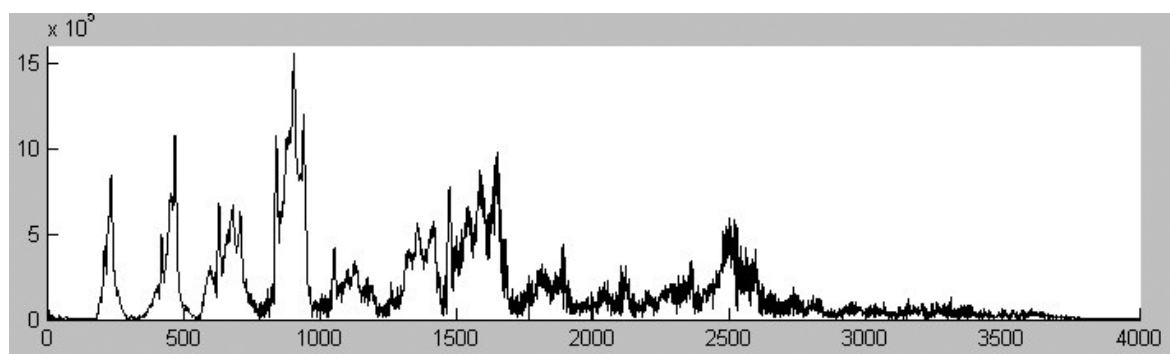
Bardzo często w celu analizy sygnału dźwiękowego korzysta się z jego częstotliwościowej reprezentacji. Mówimy wtedy o tzw. **widmie** sygnału dźwiękowego. Widmo sygnału dźwiękowego na obrazowanie jakie składowe sinusoidalne będące funkcjami czasu (o jakich częstotliwościach i amplitudach) tworzą dany dźwięk. Na rys. 3 jest przedstawione przykładowe widmo sygnału dźwiękowego. Na osi x przedstawione są częstotliwości składowych sinusoidalnych. W tym przypadku zawierają się one w zakresie 0Hz – 4000Hz. Z osi y można odczytać pośrednio informację o amplitudach składowych sinusoidalnych.



Rysunek 2.

Sygnał sinusoidalny

Układy przetwarzania czy odtwarzania dźwięku często w pewien sposób zniekształcają sygnał, wpływają na zawartość składowych sinusoidalnych w sygnale. Najczęściej nie jest to zjawisko pożądane. Dla układów fonicznych określa się charakterystyki częstotliwościowe – amplitudową i fazową, które są jednocześnie wyznacznikiem wprowadzanych przez układ tzw. zniekształceń liniowych. Charakterystyka amplitudowa powinna być maksymalnie płaska w funkcji częstotliwości (w paśmie układu). Natomiast charakterystyka fazowa powinna być maksymalnie liniowa w funkcji częstotliwości (również w paśmie układu). Zniekształcenia liniowe najczęściej można w pewien sposób skorygować metodami analogowymi bądź cyfrowymi. Oprócz tego w układach fonicznych pojawiają się też zniekształcenia nieliniowe (np. po przejściu przez wzmacniacz akustyczny sygnału dźwiękowego mogą pojawić się składowe sinusoidalne, których wcześniej w sygnale w ogóle nie było). Takie zniekształcenia są trudne do skorygowania.



Rysunek 3.

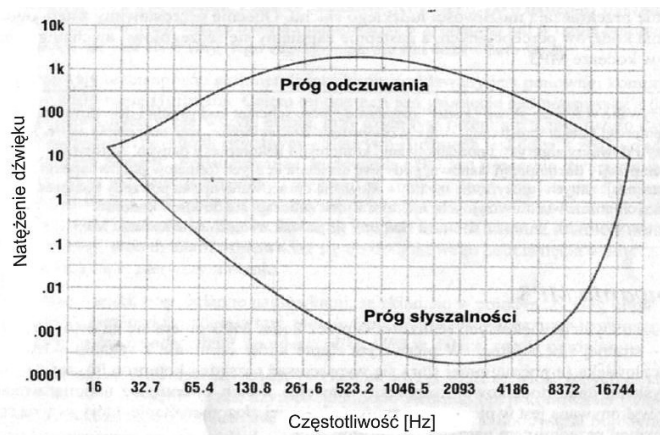
Widmo sygnału dźwiękowego

1.2.1 ZAKRES SŁYSZALNOŚCI

Głośność to taka cecha wrażenia słuchowego, która umożliwia uszeregowanie dźwięków na skali głośności. Teoretycznie ucho ludzkie potrafi odebrać i przetworzyć drgania o częstotliwości 16Hz do 20kHz. Jest to jednak duże uproszczenie nie mające wiele wspólnego z rzeczywistością. Okazuje się, że powyższy zakres jest słyszalny tylko wtedy, gdy energia dźwięku jest duża. Przy cichych dźwiękach czułość ucha drastycznie maleje w obszarze częstotliwości poniżej 200Hz oraz powyżej 8kHz. W tych zakresach trudniej jest również rozróżniać wysokość dźwięku. Zakres częstotliwościowy percepcji dźwięków maleje też wraz z wiekiem.

Na wrażenie głośności dźwięku wpływa wiele dodatkowych czynników. Po pierwsze wrażenia głośności zależą od czasu trwania dźwięku. Dla krótkich czasów trwania dźwięków występuje efekt czasowego sumowania głośności. Natomiast dla czasów od ok. 1 sekundy do ok. 3 minut dla dźwięków o niskim poziomie lub wysokiej częstotliwości głośność maleje ze wzrostem czasu trwania. Jest to efektem adaptacji głośności. W wyniku efektu sumowania głośności powiększenie szerokości pasma częstotliwościowego szumu białego powoduje wzrost głośności. Głośność szumu (i dźwięków złożonych) jest wyższa niż tonów (sinusoidalnych) o takim samym natężeniu dźwięku.

Próg słyszalności (próg absolutny, próg detekcji sygnału) jest to najmniejszy poziom ciśnienia akustycznego dźwięku, który wywołuje zaledwie spostrzegane wrażenie słuchowe wobec braku innych dźwięków. Najniższa wartość ciśnienia akustycznego (przy częstotliwości 1000Hz) wykrywanego przez ucho ludzkie wynosi średnio 20μPa (rys. 4).



Rysunek 4.
Zakres słyszalności człowieka

Próg bólu jest to wartość ciśnienia akustycznego, przy której ucho odczuwa wrażenie bólu. Jest ono prawie niezależne od częstotliwości i wynosi 140dB dla dźwięków sinusoidalnych oraz 120dB dla szumów. Wrażenie bólu wywołane jest reakcją mięśni bębienka i kosteczki ucha środkowego na impulsy wysokiego ciśnienia akustycznego. Reakcja ta ma na celu ochronę aparatu słuchowego przed ewentualnymi uszkodzeniami.

Okazuje się, że człowiek nie słyszy jednakowo dobrze wszystkich dźwięków o tym samym poziomie głośności. Dźwięki bardzo niskie i bardzo wysokie są słyszane słabo, za to tony o częstotliwościach od 1KHz do 5KHz (mniej więcej zakres mowy ludzkiej) są słyszane wyjątkowo dobrze. Na przykład ton 10dB mający częstotliwość 1000Hz będzie przez większość ludzi świetnie słyszalny, ale ton 10dB o częstotliwości 25Hz chyba wszyscy odbierzemy jako ciszę. Uświadomienie sobie faktu, że nie wszystkie dźwięki o tej samej energii są przez ludzkie ucho rozpoznawane jako tak samo głośne, to dopiero początek problemów związanych z pojęciem głośności. Następnym problemem jest fakt, że ucho działa nieliniowo. Oznacza to, że dwa razy większe natężenie dźwięku wcale nie jest przez nas odbierane jako dwa razy głośniejszy dźwięk. Ucho dokonuje silnego spłaszczenia odczuwania głośności – dźwięk, który odczuwamy jako kilka razy głośniejszy od początkowego, ma w rzeczywistości energię dziesiątki, a nawet setki razy większą.

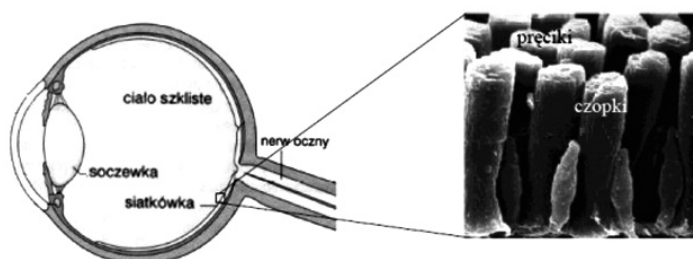
1.2.2 OCENA JAKOŚCI DŹWIĘKU

Układ słuchowy, tak jak wzrokowy, jest instrumentem nieliniowym, a odbierane przez niego dźwięki są interpretowane w sposób subiektywny. Wpływ na sklasyfikowanie odbieranego dźwięku mają między innymi wspomnienia, wiedza, doświadczenie i uszkodzenia narządu słuchowego. Subiektywna ocena jakości dźwięku przeprowadzona przez dwie osoby może dać zatem bardzo różny wynik.

1.3 OBRAZ

Zmysł wzroku odgrywa w życiu człowieka niezwykle istotną rolę, związaną nie tylko z czysto fizycznym rozpoznawaniem i rozróżnianiem otaczających nas przedmiotów i zjawisk, ale wrażenia wzrokowe wpływają także na naszą psychikę, czy nastrój. Warto również podkreślić, że tą drogą mózg człowieka przyswaja największą ilość informacji z otaczającego nas świata. Z fizycznego punktu widzenia rejestracja promieniowania świetlnego jest realizowana na siatkówce oka. Siatkówkę oka można przyrównać do pewnego rodzaju światłoczułej matrycy, na której znajdują się receptory widzenia. Takimi receptorami są **pręciki**, które rejestrują jedynie natężenie światła, bez możliwości jego analizy barwnej oraz **czopki**, które reagują na światło o określonej barwie (rys. 5). Widzenie barwne jest wynikiem reakcji fotochemicznej, w której substancje białkowe zawarte w czopkach, zwane **opsynami**, reagują na światło absorbując poszczególne składowe promieniowania barwnego. Istnieją trzy rodzaje opsyn: absorbujące światło niebieskie, zielone i czerwone. Umożliwiają one barwne widzenie dzienne. Brak opsyny jednego rodzaju (np. absorbującej światło czerwone) powoduje niezdol-

ność rozróżniania pewnych barw. W wyniku reakcji fotochemicznych energia świetlna zostaje przekształcona na impulsy nerwowe, które są dalej przesyłane przez nerw wzrokowy. Sygnały świetlne docierające do mózgu są zamieniane na cechy, takie jak: kształt, kolor, czy wzajemne relacje przestrzenne obiektów.



Rysunek 5.
Budowa oka

Obrazy cyfrowe reprezentują te same sceny, które możemy obserwować, ale przedstawione w postaci dwuwymiarowych tablic pikseli. Technika cyfrowa umożliwia przeprowadzenie wielu operacji obróbki obrazu, w tym także działań niewykonalnych tradycyjnymi metodami przy pomocy szklanych filtrów optycznych lub analogowej elektroniki. Jedną z pierwszych prób wykorzystania techniki cyfrowej w praktyce było przesyłanie obrazów na odległość z wykorzystaniem kabla. Do przesłania obrazów użyto opracowanego w 1920 r. tzw. „Bartlane System”, który umożliwiał skanowanie obrazu element po elemencie. Negatyw fotografii był poddawany naświetleniom w pięciu różnych czasach ekspozycji. Obraz był rejestrowany na płytkach cynkowych. Każdy punkt zdjęcia był tym samym charakteryzowany kombinacją pięciu bitów opisującą wzrastającą jasność obrazu. W wyniku skanowania powstawała taśma papierowa rejestrująca poziomy szarości obrazu (5 bitowe kodowanie). Wykorzystanie kabla transatlantyckiego umożliwiło przesłanie obrazów przez ocean Atlantycki (Londyn – Nowy York). Nietrudno sobie wyobrazić jak bardzo uprościło to wymianę informacji. Jeden z pierwszych przesyłanych tą drogą obrazów jest przedstawiony na rys. 6.

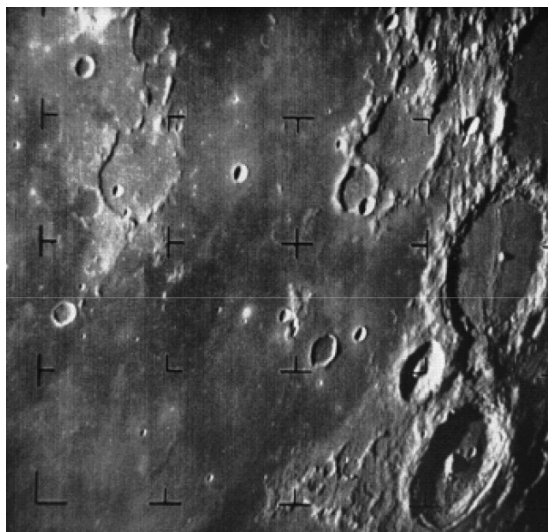


Rysunek 6.
Obraz przetransmitowany i odtworzony przez Bartlane System

Dalszy, bardzo szybki rozwój technik cyfrowych nastąpił w latach 1939-45. w czasie II Wojny Światowej bardzo potrzebne były efektywne systemy rozpoznawania wojskowego, prowadzono więc szeroko zakrojone badania w tym kierunku. Techniki cyfrowe wykorzystano głównie do podwyższania jakości obrazu fotograficznego (dystorsja, nieostrość, kontrast). Początek lat 60. XX wieku to jednocześnie początek misji kosmicznych NASA (misje Ranger'a). Na rysunku 7 przedstawiono obraz Księżyca sfotografowany przez statek Ranger 7. Zdjęcie zostało wykonane w 1964 roku przy użyciu kamery telewizyjnej i następnie przesłane na Ziemię. Zdję-



cia z tej misji uzmysłowiły konieczność intensyfikacji w rozwoju metod przetwarzania i analizy obrazu. Obecnie cyfrowe przetwarzanie obrazów jest praktycznie wszechobecne. Trudno jest znaleźć dziedzinę nauki, a także i rozrywki, w której nie byłoby jakiegoś elementu związanego z cyfrowym przetwarzaniem obrazów.



Rysunek 7.

Pierwszy obraz Księżyca sfotografowany przez statek Ranger 7

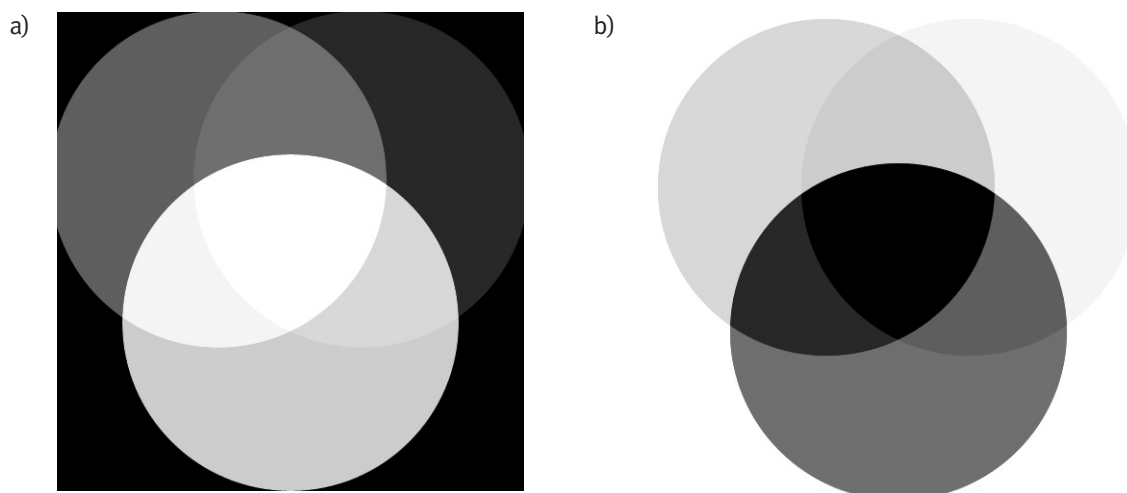


1.3.1 MODELE BARW

Barwa jest wrażeniem psychicznym wywołanym w mózgu człowieka w czasie, gdy oko rejestruje promieniowanie elektromagnetyczne z widzialnej części fal świetlnych. Główny wpływ na to wrażenie ma skład widmowy promieniowania świetlnego, ilość energii świetlnej, obecność innych barw w polu widzenia obserwatora, ale także cechy osobnicze obserwatora, takie jak: zdrowie, samopoczucie, nastrój, a nawet doświadczenie i wiedza w posługiwaniu się własnym organem wzroku. Barwa z samej swojej natury jest trudna do zdefiniowania, stąd tworzy się mnóstwo wzorców, tabel i modeli próbujących uporządkować barwy. Modele barw są próbą opisu barw przy użyciu pojęć matematycznych. Przy opisie sprzętu najczęściej wykorzystywanymi modelami barw są modele RGB i CMY/CMYK.

Model barw RGB jest modelem barw ukierunkowanym na sprzęt, w którym barwa powstaje w wyniku emisji światła: monitory, skanery, cyfrowe aparaty fotograficzne. Jest to model addytywny, w którym wszystkie barwy powstają przez zmieszanie trzech barw podstawowych: czerwonej, zielonej i niebieskiej. **Mieszanie addytywne** (rys. 8a) to mieszanie barw poprzez sumowanie wiązek światła widzialnego różnych długości. Synteza addytywna zachodzi np. podczas projekcji na ekran: w miejscu oświetlonym jednocześnie światłem o różnej barwie oko ludzkie widzi odbity strumień światła będący sumą wszystkich padających w to miejsce barw (w widzianym przez nas strumieniu odbitym występują na raz wszystkie długości fal odpowiadające poszczególnym strumieniom światła padającego).

Model barw CMY jest modelem ukierunkowanym na sprzęt drukujący: drukarki, maszyny drukarskie. Wrażenie barwy uzyskuje się dzięki światłu odbitemu od zadrukowanego podłoża. Pigment farb/atramentów pochłania określone długości fali, a odbija pozostałe. Dlatego model ten jest nazywany modelem substraktywnym. **Mieszanie substraktywne** to mieszanie barw poprzez odejmowanie wiązek światła odpowiadającego różnym długościom fal (najczęściej realizowane jest poprzez pochłanianie niektórych długości fal przez powierzchnię, od której odbija się światło białe). Synteza substraktywna zachodzi np. przy mieszanii farb o różnych barwach: w miejscu pokrytym farbą (powstałą ze zmieszania farb o różnych barwach) oko ludzkie widzi odbity strumień światła będący tą częścią światła białego, która zostanie po pochłonięciu wszystkich składowych barwnych przez poszczególne farby wchodzące w skład mieszanki (rys. 8b). Wszystkie barwy w modelu CMY powstają przez zmieszanie trzech barw podstawowych: cyan (zielono-niebieska), magenta (purpurowa), yellow (żółta). Zmieszanie C, M i Y powoduje odfiltrowanie całego światła i powstaje kolor czarny. W praktyce trudno jest uzyskać w ten sposób idealny kolor czarny. Dlatego powstał model **CMYK**, w którym zdecydowano się na dodanie jeszcze jednego koloru – czarnego (black).



Rysunek 8.
Addytywne a) i subtraktywne b) mieszanie barw

Barwy można opisać używając atrybutów barw. Atrybuty barwy to odcień, nasycenie i jasność. **Odcień** jest cechą jakościową barwy związaną z długością fali dominującej w strumieniu światła. Przy widzeniu barwnym obserwując poszczególne pasma widma o różnych długościach fali stwierdzimy, że istnieje charakterystyczna różnica między każdym z tych wrażeń. Doznawane wrażenia określamy nazywając je kolejno: fioletowy, niebieski, zielony, żółty, pomarańczowy, czerwony. Tę cechę wrażenia wzrokowego nazywamy właśnie odcieniem barwy. **Nasycenie** jest cechą jakościową barwy i podaje stosunek ilości światła monochromatycznego do ilości światła białego – im większe nasycenie, tym mniejszy jest udział w widmie promieniowania fal o innych długościach niż fali dominującej. **Jasność, jaskrawość** jest cechą ilościową barwy. Jasność dotyczy obiektów odbijających światło, jaskrawość – świejących i odpowiada wrażeniu słabszego lub mocniejszego strumienia światła.

1.4 ZASTOSOWANIE DŹWIĘKU W MULTIMEDIACH

Upowszechnienie się multimediiów w dużej mierze zawdzięczamy rozwojowi komputerów i technik komputerowych. Coraz szybsze i lepsze komputery umożliwiały łączenie różnych metod przekazu informacji, w szczególności:

- zastosowanie w komputerach kart dźwiękowych umożliwiło łatwą i jednocześnie zaawansowaną obróbkę sygnałów dźwiękowych na potrzeby prezentacji multimedialnych,
- komputerowe programy do obróbki grafiki pozwoliły na przetwarzanie statycznych obrazów dwuwymiarowych,
- standard wi-fi, w połączeniu z coraz silniejszymi obliczeniowo komputerami i dostępnością kamer cyfrowych umożliwiły również stosowanie nagrań wideo.

W dalszej części opracowania poznamy sprzęt i oprogramowanie umożliwiające posługiwanie się multimediami jak i tworzenie interesujących przekazów multimedialnych. Niewątpliwie jednymi z pierwszych i jednocześnie bardzo istotnych urządzeń tego typu były karty dźwiękowe.

1.4.1 KARTY DŹWIĘKOWE

Karty dźwiękowe są centralnym punktem każdego domowego studia nagrań. Warto więc zapoznać się dokładniej z ich budową i zasadą działania. Umożliwi to w pełni wykorzystanie możliwości współczesnych kart dźwiękowych, a są one niemałe.

Historia komputerów PC i generowanego przez nich dźwięku może nie jest zbyt długa, lecz z pewnością interesująca. Już w pierwszych komputerach PC z początku lat 80. XX wieku montowano głośniczki, które przez długie lata były jedynie namiastką tego, co może dzisiaj dokonać komputer w dziedzinie przetwarzania dźwięków. W 1986 roku pojawił się tzw. Speech Thing firmy Covox. Urządzenie to miało postać wtyczki wkładanej do portu drukarki i stanowiło jednokanałowy, ośmiobitowy przetwornik cyfrowo-analogowy. Umożliwiało jednak



odtworzenie cyfrowo zapisanych odgłosów. To dość kosztowne urządzenie składało się jedynie z kilku oporników i kondensatorów upakowanych do wtyczki od drukarki. W Polsce zaczęto masową produkcję tzw. Covoxów, czyli chałupniczych kopii Speech Thing, które szybko stały się bardzo popularne ze względu na niską cenę.

W 1987 roku firma AdLib zaprezentowała swoją kartę muzyczną do komputerów PC. Była to pierwsza karta, która naprawdę zyskała ogromną popularność. Później jej konstrukcję wykorzystano przy projektowaniu słynnego **Sound Blaster**. AdLib miała postać niewielkiej karty rozszerzenia ISA. Zastosowano w niej procesor Yamaha YM3812, który korzystał z syntezy dźwięku metodą modulacji częstotliwości (FM – *Frequency Modulation*). **Synteza FM** polega na mieszaniu paru fal różnych kształtów, otrzymujemy wówczas charakterystyczny, syntetyczny dźwięk. Syntezator karty AdLib zdolny był do generowania jednocześnie dziewięciu głosów instrumentalnych lub sześciu głosów instrumentalnych plus pięciu perkusyjnych (co dawało łącznie 11 głosów). Niestety, na ogół nie udawało się uzyskać dźwięku wiernie naśladowującego jakikolwiek instrument muzyczny. Sama synteza FM miała swą premierę na długo przed pojawieniem się kart AdLib. Od przeszło dziesięciu lat stosowano ją w branży muzycznej do muzyki elektronicznej (z syntezy FM korzystali m.in. Jean Michel Jarre oraz Mike Oldfield).

W 1989 roku firma Creative Music Systems (dziś znana jako Creative Labs) wyprodukowała kartę Sound Blaster. Karta zawierała syntezator FM, odtwarzała dźwięk mono z rozdzielczością 8 bitów. Sound Blaster można określić jako wzbogaconą o ośmiobitowy przetwornik cyfrowo-analogowy (c/a) i analogowo-cyfrowy (a/c) kartę AdLib. Główny element karty stanowił ten sam procesor firmy Yamaha, zatem możliwości muzyczne karty Sound Blaster nie zmieniły się w stosunku do karty AdLib. O jej popularności zdecydowała obecność przetworników c/a i a/c. Wreszcie możliwe było samodzielne zgrywanie cyfrowego dźwięku. Sound Blaster do swej monofonicznej muzyki potrafił też dodawać prawdziwe, cyfrowe efekty dźwiękowe. Karta zdobyła dzięki temu niezwykłą popularność.

Kolejny przełom nastąpił w 1992 roku, kiedy to znana jedynie z joysticków firma Gravis zaprezentowała pierwszą popularną, tanią kartę muzyczną WaveTable o nazwie UltraSound (zwaną GUS – *Gravis UltraSound*), z procesorem Gravis GF1. WaveTable (tablica fal) oznacza nową, rewolucyjną w tej klasie cenowej metodę generowania dźwięku, znaną wcześniej jedynie z bardzo drogich, profesjonalnych instrumentów elektronicznych.

Z uwagi na sztuczne brzmienie generowanych dźwięków, synteza FM nie nadawała się do zastosowań profesjonalnych. Z tego względu producenci sprzętu opracowali technikę **syntezy wavetable** (WT). Zasada działania syntezy WT jest bardzo prosta. W celu uzyskania brzmienia instrumentu, chip muzyczny nie generuje sztucznego dźwięku, lecz odtwarza jego oryginalny dźwięk, nagrany wcześniej w studiu. W praktyce nie ma możliwości zapisania w pamięci wszystkich dźwięków instrumentów. Chip muzyczny musi więc często obliczać wysokość i długość dźwięków na podstawie wzorcowych próbek. Z zadaniem tym poszczególne karty WT radzą sobie bardzo różnie. W niektórych modelach można np. uzyskać lepsze brzmienie instrumentów smyczkowych w innych instrumentów dętych. Naprawdę dobre brzmienie dla wszystkich odmian muzyki oferują kosztowne karty profesjonalne.

Firma Creative Labs, w krótkim czasie po premierze Gravisa, zaprezentowała własną kartę WaveTable – AWE32 z 32-głosowym procesorem EMU8000. AWE32 był pod względem budowy bardzo podobny do karty GUS, aczkolwiek nie był z nią absolutnie zgodny. Karta AWE32 umożliwiała jednak rozszerzenie pamięci na próbki dźwiękowe do bardzo dużej wówczas wielkości 28 MB.

Od 1998 roku standardem na rynku kart muzycznych stał się dźwięk przestrzenny. Wykorzystuje się go zarówno w grach jak i do kina domowego. W 1998 roku powstała karta Sound Blaster Live! – pierwsza wielokanałowa karta dźwiękowa, z własnym procesorem dźwiękowym EMU10K1 i nowym standardem dźwięku EAX.

W późniejszych wersjach kart znaczną poprawę jakości dźwięku uzyskano dzięki 24-bitowym sześciokanałowym przetwornikom cyfrowo-analogowym, generującym dźwięk o częstotliwości 96kHz. Karty z takimi przetwornikami mają stosunek sygnału do szumu (SNR) wynoszący ponad 100dB. To rekord jeśli chodzi o karty dla użytkowników domowych. W kartach firmy Creative Labs zastosowano unikatowy mechanizm symulacji pogłosu EAX, który umożliwia generowanie czterech rodzajów efektów akustycznych:

- *Environment Morphing* – umożliwia naturalne przechodzenie dźwięku z jednego środowiska do drugiego ze zniekształcaniem brzmienia w czasie rzeczywistym, zgodnie z charakterystyką akustyczną każde go środowiska.



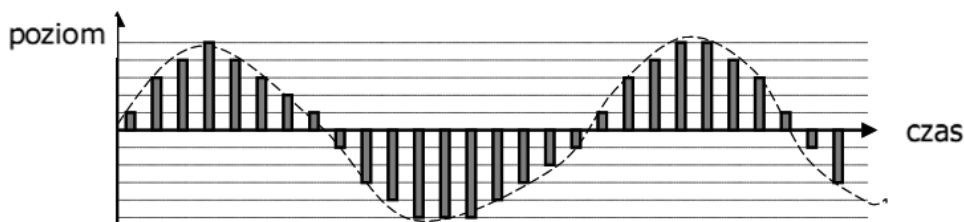
- *Panning* – umożliwia usłyszenie trójwymiarowej przestrzeni i lokalizowanie bliskich i odległych dźwięków.
- *Environment Reflections* – odtwarza naturalne zjawisko odbijania się dźwięku od powierzchni i powracania do słuchacza w pełnej przestrzeni trójwymiarowej.
- *Environment Filtering* – wykorzystuje górnoprzepustowe filtry dźwięku do dokładnej symulacji dźwięku ze środowisk wewnętrznych i zewnętrznych.

1.4.2 BUDOWA KART DŹWIĘKOWYCH

Karty dźwiękowe w zależności od stopnia złożoności i zaawansowania mogą zawierać różne elementy. Najważniejsze z nich to:

- **generator dźwięku** – występował w starszych kartach i był to zazwyczaj generator FM oraz generator szumu, służył do sprzętowego generowania dźwięków za pomocą modulacji i łączenia fal oraz szumu;
- **pamięć ROM lub półprzewodnikowa typu flash** – umożliwia przechowywanie danych (np. próbek wykorzystywanych do syntezy dźwięku);
- **przetworniki a/c i c/a** – umożliwiające rejestrację i odtwarzanie dźwięku;
- **mikser dźwięku** – służy do łączenia sygnałów dźwięku z różnych źródeł, generatorów dźwięku, przetworników c/a, wejść zewnętrznych, itp.;
- **wzmacniacz wyjściowy** – służy do podłączenia słuchawek lub dopasowania linii wyjściowych przetwornika c/a;
- **interfejs do komputera** – służy do komunikacji i wymiany danych z kartą dźwiękową, zazwyczaj ISA, PCI lub USB;
- **procesor DSP** – służy do cyfrowej obróbki dźwięku, np. nakładania efektów;
- **interfejs MIDI** – służy do podłączania do komputera cyfrowych instrumentów muzycznych.

Jednym z najważniejszych elementów karty dźwiękowej jest **przetwornik analogowo-cyfrowy a/c** (ang. A/D – *analog to Digital* lub ADC – *analog to digital converter*). Układ ten służy do zamiany sygnału analogowego (ciągłego) pochodzącego od obiektów świata realnego na reprezentację cyfrową (sygnał cyfrowy). Przetwarzanie a/c składa się z dwóch etapów: próbkowania i kwantyzacji. Analogowy sygnał jest ciągły w czasie, więc konieczne jest przetworzenie go na ciąg liczb (rys. 9).



Rysunek 9.

Reprezentacja cyfrowa sygnału analogowego

To, jak często sygnał jest sprawdzany i, zależnie od jego poziomu, zamieniany na liczbę, jest określane mianem **częstotliwości próbkowania**. Innymi słowy można powiedzieć, że częstotliwość próbkowania jest odwrotnością różnicy czasu pomiędzy dwoma kolejnymi próbkami. Wiarygodne odwzorowanie sygnału jest możliwe do osiągnięcia, gdy częstotliwość próbkowania jest większa niż podwojona, najwyższa składowa częstotliwość sygnału przetwarzanego (twierdzenie Nyquista-Shannona). **Kwantowanie** polega na zastąpieniu wartości zmieniających się płynnie wartościami zmieniającymi się skokowo w odpowiedniej skali (dokładności) odwzorowania. Różnica pomiędzy wartością sygnału analogowego i skwantowanego to tzw. błąd kwantyzacji. **Rozdzielczość przetwornika** określa liczbę dyskretnych wartości jakie może on wytworzyć. Zwykle wyraża się ją w bitach. Przykładowo, przetwornik a/c, który potrafi przetworzyć próbkę sygnału na jedną z 256 wartości liczbowych ma rozdzielczość 8 bitów, ponieważ $2^8 = 256$. Rozdzielczość może być również wyrażona w voltach. Rozdzielczość napięcia przetwornika a/c jest równa jego całkowitej skali pomiaru podzielonej przez liczbę poziomów kwantyzacji.

Innym istotnym elementem karty dźwiękowej jest **procesor sygnałowy** (ang. *Digital Signal Processor*, procesor DSP). Jest to specjalizowany procesor do cyfrowej obróbki sygnałów. Procesory takie charakteryzują się

rozdzielonymi pamięciami programu i danych (architektura harwardzka), możliwością równoczesnego odczytu instrukcji i danych, sprzętowym dostosowaniem do wykonywania operacji najczęściej występujących przy przetwarzaniu sygnałów (filtracji FIR i IIR, transformacji Fouriera, korelacji wzajemnej) i potokowym przetwarzaniem instrukcji.

Wyjaśnienia wymaga też standard MIDI. **Musical Instrument Digital Interface (MIDI)** jest systemem (interfejs, oprogramowanie i zestaw komend) służącym do przekazywania informacji pomiędzy elektronicznymi instrumentami muzycznymi. MIDI umożliwia wymianę informacji między komputerem, syntezatorem, kartami dźwiękowymi i podobnymi urządzeniami. Standard MIDI został utworzony w 1983 roku w celu ujednoczenia cyfrowych syntezatorów. Wraz z rozwojem komputerów osobistych i technologii multimedialnych standard MIDI został zaadaptowany do komunikacji między komputerem a kartą dźwiękową. Umożliwiło to komponowanie muzyki oraz jej odtwarzanie tylko przy użyciu komputera z odpowiednim oprogramowaniem. MIDI definiuje standardy sprzętowe oraz język komend. Informacje przesyłane są za pomocą połączenia szeregowego. Przekazywana jest standardowa informacja składająca się z takich składowych, jak: wysokość dźwięku, natężenie dźwięku, modulacja itp. jednocześnie dla 16 kanałów. W ostatnim czasie szeregowy standard MIDI został zastąpiony przez MIDI USB, gdzie do komunikacji jest wykorzystywana magistrala USB.

Obecnie na rynku PC zewnętrzne karty muzyczne straciły dużo na znaczeniu. Dźwięk 4,5 kanałowy jest już standardem w kartach muzycznych zintegrowanych ze współczesnymi płytami głównymi. Pojawiają się w nich również procesory DSP do efektów 3D. Jakość wbudowanych kart dźwiękowych ciągle rośnie zadowalając swoimi możliwościami większość nabywców, dodatkowy wydatek na naprawdę dobrą kartę muzyczną PCI, czy USB staje się zatem zbędny.

1.5 PREZENTACJA MULTIMEDIALNA

Trudno jest podać gotowy przepis na doskonałą prezentację multimedialną. Istnieje jednak kilka elementów, na które trzeba koniecznie zwrócić uwagę.

Każdą prezentację należy zacząć od przygotowania planu, albo przynajmniej należy określić główny cel swojej prezentacji. Należy zastanowić się, co prezentacja ma przekazywać, co ma być w niej najważniejsze, a co jest na tyle nieistotne, że można to pominąć. W tym miejscu nie trzeba nawet włączać komputera. Wystarczy kartka i długopis. Trzeba także zastanowić się, do kogo prezentacja ma być skierowana. Prezentacja ma być w pełni zrozumiała dla osób, które ją poznają. Trzeba więc rozważyć, co już teraz wiedzą twoi słuchacze, a co powinni dopiero poznać.

Jako że większość ludzi może zapoznać się tylko z małą liczbą faktów podczas jednej prezentacji, warto sporządzić listę tego, co ma zrozumieć publiczność po zakończeniu prezentacji. Limit „głównych myśli” powinien wynosić nie więcej niż 5-7. W przeciwnym wypadku prezentacja okaże się prawdopodobnie zbyt złożona dla większości osób. Co ważne, trzeba również rozważyć czas, jaki będzie przeznaczony na przeprowadzenie prezentacji. Ponadto należy pamiętać, aby angażować publiczność. Na slajdach w prezentacji nie może być zbyt dużo tekstu. Przekazując informację należy również stosować różnego typu grafiki, ilustracje, wykresy. Co ważne, tekstu nie może też zabraknąć całkowicie, gdyż wówczas prezentacja może wydać się zbyt prosta i poniekąd nieciekawa.

Należy dobrać taki rozmiar czcionki, aby zapewnił dobre odczytanie zawartych informacji. Najlepiej stosować czcionki proste tzn. Arial, Times New Roman, Georgia, Verdana o rozmiarze 24-28 punktów. Są one bardzo czytelne i nie męczą oka. Ważny jest również dobór kolorów. Należy unikać łączenia ze sobą kolorów z przeciwstawnymi granic widma (rys.10). Takie połączenia kolorów są źle odbierane i męczące dla wzroku. Przykład źle dobranych kolorów pokazany jest na rys. 11.



Rysunek 10.
Widmo światła widzialnego

Czerwień i niebieski znajdują się na przeciwstawnych krańcach widma. Oczom jest trudno równocześnie zogniskować się na obydwu barwach

Rysunek 11.

Źle odbierane połączenia kolorów bardzo męczą wzrok

Prezentację bardzo dobrze urozmaicają dźwięki, obrazy i filmy. Jeżeli przedstawiamy jakieś dane dobrze jest ilustrować je tabelarycznie bądź w postaci wykresów, rycin, schematów i diagramów. Można stosować animacje: obiektów na slajdach, przejść między slajdami, itp. Kolejność animacji powinna być zgodna z treścią. Animację powinny być jednak stosowane z umiarem. Będą wtedy ożywiać prezentację i zwracać uwagę oglądającego na istotne kwestie. Natomiast jeśli jest ich zbyt dużo, mogą rozpraszać uwagę.

Następny problem to rozmieszczenie elementów na slajdzie. Powinno się raczej umieszczać grafikę z boku. Mamy wtedy więcej miejsca dla tekstu. Poza tym ustawienie grafiki z lewej strony kieruje wzrok na tekst.

Również tempo prezentacji musi być dobrze przemyślane. Należy dostosować tempo wyświetlania slajdów do treści, którą widz musi przeczytać. Tempo nie może być za szybkie, ponieważ słuchacz nic nie zapamięta. Tempo nie może być również za wolne ponieważ znużymy słuchacza oraz go zniechęcimy do obejrzenia i wysłuchania dalszej części prezentacji.

Oczywiście prezentacja nie może zawierać błędów merytorycznych. Powinna zawierać informacje sprawdzone, dokładne i aktualne i w żadnym razie nie może zawierać błędów ortograficznych, stylistycznych i gramatycznych.

Zadanie 1. Utworzenie prezentacji multimedialnej

Do wykonania zadania potrzebne będą programy:

- Adobe Reader,
- Microsoft Power Point 2007,
- Open Office Impress 3.1,
- oraz komputer i rzutnik multimedialny.

Zadanie polega na utworzeniu prezentacji multimedialnej w programie Microsoft Power Point 2007 i/lub Open Office Impress. Następnie prezentację należy zapisać w pliku PDF i odczytać w programie Adobe Reader. Wykonanie zadania umożliwi poznanie zasad tworzenia dobrej prezentacji, technicznych możliwości wyżej wymienionych programów, a także ich porównania. Na koniec prezentacja powinna być zaprezentowana przez jej twórcę przy wykorzystaniu rzutnika multimedialnego.

Zasady tworzenia dobrej prezentacji są niezależne od programu, w którym się ją tworzy. Poza tym w ostatnim czasie obserwujemy ujednoczenie interfejsów obsługi programów. Dotyczy to również programów do prezentacji. Open Office Impress 3.1 jest w zasadzie zgodny z Microsoft Power Point 2003, jeśli weźmiemy pod uwagę interfejs. Również zakres możliwości i dostępnych funkcji jest bardzo podobny. Natomiast w wersji Microsoft Power Point 2007 wprowadzono nowy typ interfejsu z użytkownikiem, zawierający tzw. „wstążki”. Program zapewnia te same opcje, ale dostęp do nich jest inny. Który z interfejsów jest lepszy – najlepiej sprawdzić samemu.

Prezentacja w Power Point 2007

Program PowerPoint 2007 zawiera sporo ciekawych i profesjonalnie wyglądających szablonów. Część z nich jest wbudowana w program, natomiast zdecydowaną większość można pobrać ze stron internetowych Microsoftu. Aby wybrać szablon, należy kliknąć w menu Office (lewy górny róg) i wybrać Nowy. Na ekranie pojawi się okno o nazwie Nowa prezentacja. Po lewej stronie znajdują się szablony, podzielone na 5 kategorii. Z kolei poniżej umieszczono podmenu zatytułowane Microsoft Office Online. Przy jego użyciu można pobrać poszczególne elementy do swoich prezentacji.

Kiedy szablon jest już wybrany można dodawać kolejne slajdy i umieszczać na nich potrzebne elementy. Kolejnym krokiem powinno być formatowanie dokumentu. W tym celu można wykorzystać



Motywy (zakładka Projekt w PowerPoint, a następnie z grupy Motywy trzeba wybrać ten, który się spodoba). Po wybraniu motywu, można dostosować: Kolory, Czcionki oraz Efekty. Opcje te pomogą w sformatowaniu prezentacji według własnych wymagań.

Tabelki, diagramy, dźwięk, obrazy – wszystkie te opcje są dostępne w zakładce Wstawianie. Niektóre informacje warto przekazać inaczej, aniżeli za pomocą słów. Może się bowiem okazać, że inny sposób będzie skuteczniejszy i łatwiejszy, jeśli chodzi o zapamiętanie poszczególnych wiadomości przez słuchaczy. Zacznijmy od **wykresów**. Dane liczbowe można wprawdzie przedstawić w zwyczajnej tabeli, ale często lepiej pokusić się o nieco więcej wysiłku i zaprezentować je z użyciem dodatkowej grafiki. Taka metoda okazuje się niejednokrotnie lepszą dla wielu odbiorców, zwłaszcza podczas porównań poszczególnych danych. Podczas tworzenia wykresu, należy upewnić się, że będzie on dla każdego łatwy w odbiorze. Z tego powodu nie można zapomnieć o nadaniu mu tytułu, legendy informującej o stosownych elementach (czasami legenda jest niepotrzebna), jak i nazwach dla osi X oraz Y. Jeśli wykres będzie zawierał sporo danych, to nie oczekujemy, że słuchacze od razu dostrzegą te, które według autora są najistotniejsze i mają największą wartość. Dlatego dane, na których powinien skupić się słuchacz, muszą się wyróżniać – może to być np. inny kolor, pogrubienie, obwódka.

Niektóre dane najlepiej prezentują się w **tabeli**. Za pomocą kolumn i wierszy uda się przedstawić informacje, które inną metodą mogłyby być bardzo trudne do przyswojenia. Większość osób jest obeznana z tabelami, więc łatwo będzie im zlokalizować umieszczone w nich dane. Tutaj także obowiązuje zasada „uwypuklania” najważniejszych informacji. Typowe zastosowanie dla tabel to porównywanie określonej liczby produktów pod względem poszczególnych cech – przykładowo specyfikacji kart graficznych czy procesorów. Tabele pozwalają też wymienić wszystkie dane i porównać je w tym samym wierszu. To bardzo skuteczna i sprawdzona metoda.

Diagramy czy **obrazki** również mogą stanowić część pokazu slajdów. Trzeba się jednak zastanowić, czy będą one najlepsze do przekazania określonej informacji. Nowość w programie PowerPoint 2007, SmartArt, umożliwia tworzenie świetnie wyglądających diagramów z minimalnym wręcz nakładem pracy. Opcję SmartArt można znaleźć w zakładce Wstawianie. Funkcja ta oferuje różnego rodzaju grafiki, umieszczone w 7 kategoriach. Naprawdę warto korzystać z tej funkcji. W prezentacji można także zamieszczać zdjęcia. Są one dobrym środkiem do pokazania danego produktu czy serwisu.

Program PowerPoint zawiera również funkcję, która umożliwia umieszczanie krótkich **klipów filmowych** w prezentacjach. Można je później uruchomić w trakcie pokazu slajdów. Przy użyciu krótkich filmów łatwo jest przedstawić jakiś proces, który byłoby trudno omówić za pomocą słów. Dla przykładu, jeśli chcemy kogoś nauczyć, jak używać nowego oprogramowania, najlepiej będzie przygotować krótki film demonstrujący krok po kroku, jak korzystać z wybranych funkcji. Filmy nie powinny być dłuższe niż 2-3 min. Długie wideo mogą w końcu zmęczyć i znudzić oglądającego. W programie PowerPoint bez trudu można zamieścić także klipy audio. Nagrania dźwiękowe są zazwyczaj zastępowane filmowymi, ale czasami może się okazać, że jednak dobrze jest z nich skorzystać. Niekiedy zdjęcie bardzo ładnie prezentuje się, gdy w tle gra akurat stwarzająca odpowiedni klimat muzyka; czasem wystarczy dobry komentarz.

Animacje

Najkrótsza zasada związana z animacjami poszczególnych slajdów, napisów (np. różnego rodzaju najazdu tekstu), zdjęć i innych elementów brzmi: nie używaj ich zbyt często. Efektowne animacje były i są nadal na tyle popularne i nagminnie wykorzystywane, że zdążyły się już znudzić słuchaczy i zniechęcają ich do oglądania pokazu slajdów, niepotrzebnie wydłużając czas prezentacji. Jeśli jest stosowana animacja przy przechodzeniu ze slajdu na slajd, to należy stosować zawsze tę samą – tym sposobem nadając prezentacji smak profesjonalizmu. Nie należy używać różnych efektów tylko dlatego, że jest to możliwe. Co za dużo, to niezdrowo. Tym niemniej, czasami animacje mogą okazać się bardzo pomocne. Dla przykładu, kiedy chcesz przekazać kilka ważnych informacji na jednym slajdzie. Wówczas warto skorzystać np. z efektu zanikania – wszystkie ważne informacje pojawiają się wtedy jedna po drugiej.

Prezentacja w Open Office Impress 3.1

Open Office Impress jest produktem w niczym nie ustępującym konkurencji, czasem wręcz ją przewyższającym. O jego atrakcyjności stanowią również możliwości tworzenia naprawdę zaawansowa-



nych, w pełni interaktywnych i multimedialnych prezentacji. W tym opisie podane zostaną jedynie niektóre opcje programu Open Office Impress, służące podnoszeniu atrakcyjności wizualnej prezentacji, nadaniu jej multimedialnego charakteru.

Prezentację najłatwiej stworzyć korzystając z szablonu. Każdy szablon można w dużym stopniu modyfikować. Jest to i ciekawa i przydatna funkcja. W następnym kroku powinniśmy wprowadzać treść prezentacji. Na tym etapie dodajemy slajdy i potrzebne elementy tych slajdów. Każdy z elementów można w szerokim stopniu modyfikować. Założmy, że chcemy stworzyć nagłówek. Możemy go stworzyć jako zwykły tekst z odpowiednimi atrybutami bądź za pomocą szablonu Fontwork. Szablon Fontwork umożliwia łatwe wprowadzanie tekstu trójwymiarowego. Wybieramy go korzystając z dolnego paska narzędziowego (kwadrat z literą a w środku). Każdy Fontwork ma kilka opcji, które mogą go zmodyfikować w zasadniczy sposób, np. Fontwork-kształt to opcja, za pomocą której ustawiamy ostateczny kształt napisu.

Podobnie jak w przypadku programu Power Point 2007, naszą prezentację możemy łatwo wzbogacić o dodatkowe elementy multimedialne. Może to być zwyczajna muzyka, albo krótki filmik. Oba elementy wstawiamy, klikając: *Wstaw* ® *Film i dźwięk*. Wstawianie tabeli, wykresów czy arkuszy kalkulacyjnych jest również proste. Obiekty te możemy w pewnym stopniu modyfikować. Dwukrotne szybkie kliknięcie lewym przyciskiem myszki przeniesie nas do trybu edycji obiektu. Możemy dokonywać daleko idących modyfikacji, możliwe jest również wprowadzenie danych i ewentualna zmiana typu obiektu, np. wykresu.

Jeżeli mamy zamiar przeprowadzić prezentację na komputerze mającym dostęp do Internetu, możemy umieścić jakieś linki internetowe i w sposób estetyczny zaprezentować je odbiorcom. Dodatkowymi elementami będą z całą pewnością znajdujące się w Galerii (*Narzędzia* ® *Galeria*) tła, znaki wypunktowania i wiele innych elementów.

Wstawienie jakiejś formy animacji jest również bardzo proste w Open Office Impress. Wybieramy wtedy opcję animacja niestandardowa. Z animacją niestandardową związany jest prawy obszar roboczy, w którym istnieje przyporządkowana mu zakładka. Początkowo aktywne są tylko przyciski dodaj i pokaz slajdów. Po kliknięciu na dodaj ukazuje się okno dialogowe z wyborem efektów animacji. Znajdują się w nim cztery podstawowe zakładki. Pierwsza to zakładka Wejście, która jest związana z efektami wprowadzającymi określony obiekt na ekran. Znajdują się w niej cztery grupy efektów (podstawowe, wyraziste, umiarkowane, specjalne). Ich zastosowanie zależy głównie od naszego gustu. Zadaniem zakładki wyróżnienie jest dodanie efektów mających na celu np. podkreślenie, zaakcentowanie jakiejś treści. Z kolei zakładka Wyjście odpowiada za efekty kojarzące się z zamykaniem obiektu. Ostatnia zakładka to ścieżki ruchu. Jej zadaniem jest określenie, po jakiej drodze będzie się poruszał animowany obiekt. Przy wszystkich rodzajach efektów wybieramy szybkość ich przebiegu (powoli, średnio, szybko, bardzo szybko) oraz to, czy chcemy od razu podglądać określony efekt. Niezwykle ciekawa jest opcja chronometraż, która występuje również jako jedna z trzech zakładek w opcjach efektu animacji. W chronometrażu znajdują się ustawienia dotyczące m. in. szybkości oraz czasu rozpoczęcia animacji. Przy użyciu tej opcji można stworzyć pokaz slajdów (lub jego część), który będzie wyświetlał się sam, bez ingerencji osoby prezentującej pokaz. Opcja Animacja niestandardowa umożliwia również ustawienie różnych przejść między kolejnymi slajdami.

Słuchacze powinni:

- zastanowić co będzie celem prezentacji i kto będzie jej odbiorcą,
- wykonać prezentację w programie Power Point 2007 i/lub Open Office Impress.

Na początku należy stworzyć prezentację z wykorzystaniem szablonu – jest to sposób łatwy i prosty, ale nie zawsze daje efekt taki jaki byśmy chcieli. Dlatego w następnym kroku należy zaprojektować własny wzorzec slajdu, bądź zmodyfikować istniejący. Należy zwrócić uwagę na tło, kolorystykę, proporcje i rodzaj czcionek. Wzorzec slajdu ma stopkę, nagłówek, pola: robocze, tytułu, daty. Należy je uzupełnić i zaobserwować co zmienia się w wyglądzie slajdów. W ten sposób można np. na dole slajdu (stopka) wyświetlić numer strony.

Wykonana prezentacja może dotyczyć dowolnego tematu, ale powinna zawierać:

- stronę tytułową: tytuł prezentacji, imię i nazwisko autora,
- stronę zawierającą wyliczanie i numerowanie wielopoziomowe,
- stronę z wstawionym obrazem statycznym (grafiką),
- stronę z łączącą grafikę i tekst,



- stronę z listą animowaną (poszczególne elementy wjeżdżają po kliknięciu myszką),
- stronę zawierającą elementy grupowane (np. prostokąt, tekst, wzór matematyczny połączone w jeden obiekt), grupowanie i rozgrupowywanie obiektów,
- stronę, na której dwa elementy przesłaniają się, zachodzą na siebie (należy określić, który element jest widoczny),
- stronę z wykresem (słupkowy, kołowy, itp.), wykres powinien być odpowiednio opisany,
- animacje niestandardowe (wjazdy, podkreślenia, chronometraż), obiekty 3D.

Po przygotowaniu prezentacji należy ją zapisać w formacie charakterystycznym dla danego programu, ale również w formacie PDF.

Format PDF (ang. *Portable Document Format*, przenośny format dokumentu) jest formatem plików służącym do prezentacji, przenoszenia i drukowania treści tekstowo-graficznych. Format PDF został stworzony i później szeroko promowany przez firmę Adobe Systems. Chociaż czytnik tego formatu (program Acrobat Reader) jest udostępniany za darmo przez firmę Adobe, to oprogramowanie służące do modyfikowania tych plików jest rozpowszechniane płatnie. Istnieje jednak wiele aplikacji płatnych i darmowych, w różnym stopniu mogących wyświetlać, tworzyć i modyfikować pliki w tym formacie (np. GhostScript). Wiele programów o innym przeznaczeniu dysponuje funkcją eksportu swoich dokumentów do formatu PDF.

Istnieje możliwość szyfrowania dokumentów PDF i specyfikowania ograniczeń czynności, które nie mogą być dokonywane przez użytkownika. Żeby otworzyć zaszyfrowany dokument należy znać hasło użytkownika. Pierwotnie stosowany był szyfr o 40-bitowym kluczu, co nie zapewniało jednak wystarczającego bezpieczeństwa. W nowszych wersjach używa się szyfru o kluczu 128-bitowym, co czyni go znacznie odporniejszym na próby złamania.

Format PDF powstał jako format wynikowy, mający zachować pełny wygląd dokumentu po wydrukowaniu. Dokumenty zapisane jako PDF czytane są na ekranie monitora stwarzają wiele problemów użyteczności: w darmowym czytniku dla systemu Windows – Adobe Reader nie ma możliwości dodania zakładki; trudności napotyka skopiowanie tekstu z dokumentu. W przypadku tekstów polskojęzycznych traci się niekiedy polskie znaki.

Słuchacze powinni zapisać swoją prezentację w formacie PDF i odczytać w programie Adobe Reader. Wcześniej należy zapoznać się z możliwościami programu. Czy są one takie same, jak mają programy Microsoft Office Power Point 2007 i Open Office Impress? Czy prezentacja zapisana w PDF i odtworzona w Adobe Reader zachowała wszystkie elementy?

Na końcu słuchacze powinni zaprezentować swoje prezentacje przy wykorzystaniu rzutnika multimedialnego. Przy tej okazji można porównać możliwości programów Microsoft Office Power Point 2007 i Open Office Impress i Adobe Reader.

2. KODOWANIE I PRZETWARZANIE SYGNAŁÓW DŹWIĘKOWYCH

2.1 REJESTRACJA SYGNAŁÓW DŹWIĘKOWYCH

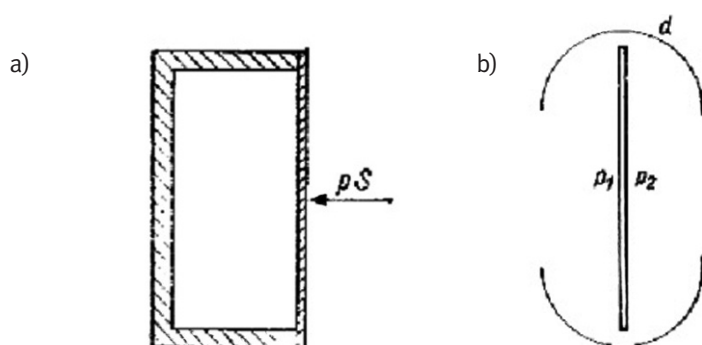
Aby możliwa była reprodukcja dźwięku, w większości przypadków jest konieczna najpierw rejestracja dźwięku. Do tego celu stosuje się różnego typu mikrofony, połączone z urządzeniami odbierającymi od mikrofonu sygnały elektryczne i rejestrującymi je. Mikrofon jest przetwornikiem elektroakustycznym, przetwarzającym sygnały akustyczne w sygnały elektryczne. Mikrofony można podzielić ze względu na rodzaj przetwornika (podział elektryczny): dynamiczne, pojemnościowe oraz ze względu na oddziaływanie pola akustycznego na membranę (podział akustyczny): ciśnieniowe, gradientowe, ciśnieniowo-gradientowe. Każdy mikrofon składa się z membrany odbierającej falę dźwiękową i przetwornika drgań membrany na zmiany wielkości elektrycznej, np. rezystancji, pojemności, napięcia. Najważniejszymi parametrami mikrofonów są: zakres przetwarzanych częstotliwości, czułość (zwana również skutecznością), charakterystyka kierunkowości.

2.1.1 PODZIAŁ AKUSTYCZNY MIKROFONÓW

Mikrofony ciśnieniowe są wszechkierunkowe. Membrana w tych mikrofonach jest otwarta tylko z jednej strony (rys. 12a). Zasada ich działania jest oparta na pomiarze ciśnienia akustycznego. **Mikrofony gradientowe**



mają charakterystykę dwukierunkową (ósemkową). W tych mikrofonach mierzony jest gradient (różnica) ciśnień pomiędzy dwiema stronami membrany (rys. 12b).



Rysunek 12.

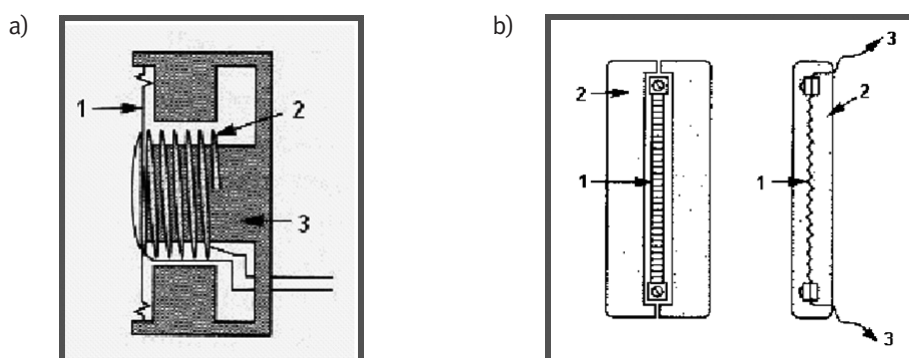
Zasada działania mikrofonu ciśnieniowego a) i gradientowego b)

Mikrofony ciśnieniowo-gradientowe są połączeniem mikrofonu gradientowego i ciśnieniowego. Ich charakterystyka kierunkowa zawiera się pomiędzy ósemką a kołem. Szczególnie popularną grupą tego typu mikrofonów są mikrofony o charakterystyce kardoidalnej.

2.1.2 PODZIAŁ MIKROFONÓW ZE WZGLĘDU NA RODZAJ PRZETWORNIKA (PODZIAŁ ELEKTRYCZNY)

W mikrofonach magnetoelektrycznych (dynamicznych) wykorzystuje się zjawisko indukcji magnetycznej. Do membrany jest przyklejona lekka cewka umieszczona w szczelinie magnesu trwałego. Podczas drgań membrany mikrofonu, w uzwojeniu cewki indukuje się siła elektromotoryczna o częstotliwości równej częstotliwości drgań membrany. Mikrofony dynamiczne można podzielić na:

- **cewkowe** – fala akustyczna oddziałuje na membranę połączoną z cewką, umieszczoną w polu magnesu trwałego. Drgania cewki powodują powstanie zmiennego sygnału elektrycznego proporcjonalnego do zmian padającej fali (rys. 13a).
- **wstęgowe** – padająca fala akustyczna pobudza do drgań wstęgę umieszczoną w polu magnetycznym, co powoduje powstanie proporcjonalnego do niej sygnału elektrycznego (rys. 13b).



Rysunek 13.

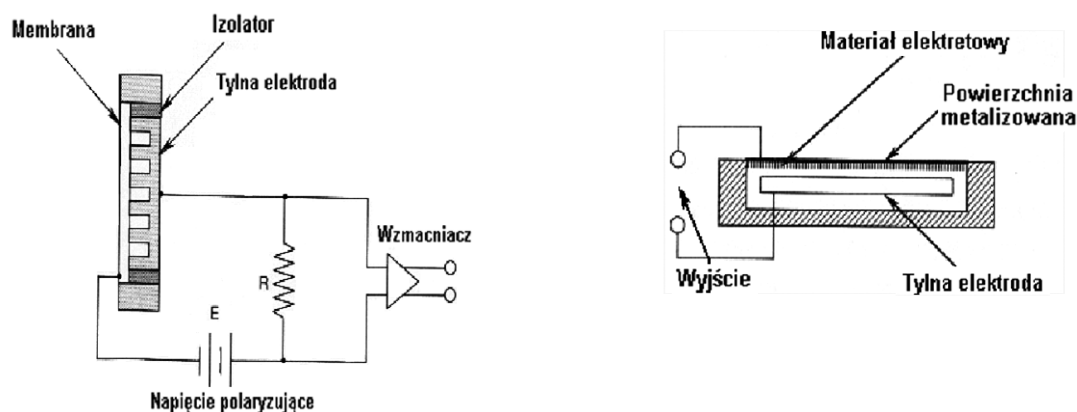
Mikrofony dynamiczne: a) cewkowe 1 – membrana 2 – cewka 3 – magnes;

b) wstęgowe 1 – wstęga 2 – magnes 3 – transformator

W przypadku **mikrofonu pojemnościowego** membrana oraz tylna elektroda są naładowane elektrostatycznie. Drgania membrany powodują zmiany napięcia pomiędzy nią i drugą elektrodą, proporcjonalnie do padającej fali. W zależności od sposobu polaryzacji elektrod wyróżniamy mikrofony: z polaryzacją zewnętrzną oraz z polaryzacją wewnętrzną (elektretowe).

Mikrofony pojemnościowe (rys. 14a) mają małe wymiary i masę, prostą budowę i małą wrażliwość na wstrząsy. Wychylenia membrany tych mikrofonów są bardzo małe – dochodzą najwyżej do kilku tysięcz-

nych części mikrometra. Mikrofon pojemnościowy, jako przetwornik dźwięku na napięcie, wymaga spolaryzowania napięciem wstępnym. Bardzo wygodne w stosowaniu są mikrofony pojemnościowe z polaryzacją wewnętrzną – tzw. **mikrofony elektretowe**. Membrana w mikrofonach elektretowych jest wykonana z folii elektretowej pokrytej od strony zewnętrznej warstwą metalu (jest to jedna okładzina naładowanego kondensatora). Drugą okładziną kondensatora jest sztywna płytki (rys. 14b). Mikrofony pojemnościowe są powszechnie stosowane w przenośnych urządzeniach do nagrywania dźwięków, np. w magnetofonach, videokamerach itp.



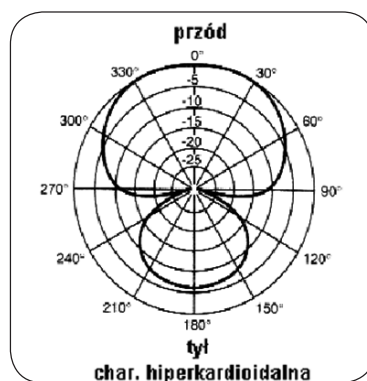
Rysunek 14.

Zasada działania mikrofonu pojemnościowego a) i elektretowego b)

2.1.3 PARAMETRY MIKROFONÓW

Parametry mikrofonów zmieniają się wraz z częstotliwością. Dlatego większość z nich jest podawana w postaci charakterystyk częstotliwościowych. Najważniejsze z nich to: charakterystyka kierunkowa, charakterystyka widmowa, czułość, impedancja.

Charakterystyka kierunkowa określa obszar, z jakiego zbierany jest sygnał o jednakowym poziomie (rys. 15). Wyróżnia się pięć podstawowych charakterystyk: dookólną (wszechkierunkową), ósemkową (dwukierunkową), kardioidalną, superkardioidalną oraz hiperkardioidalną. Charakterystyka kierunkowości mikrofonu zależy głównie od konstrukcji jego obudowy. Mikrofony o silnej kierunkowości są wykorzystywane np. przy nagrywaniu wywiadów w miejscach publicznych, nagrywaniu śpiewu ptaków itp.



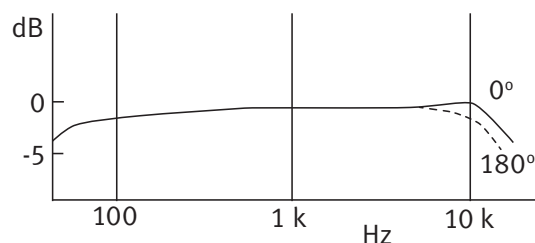
Rysunek 15.

Przykładowa hiperkardioidalna charakterystyka kierunkowa mikrofonu

Z kolei **charakterystyka widmowa** określa poziom sygnału na wyjściu mikrofonu w funkcji częstotliwości (rys.16). Najbardziej wierne przetwarzanie sygnału uzyskuje się w przypadku płaskiej charakterystyki widmowej. Często jednak mikrofony są dedykowane do przetwarzania konkretnych dźwięków, np. mowy, instrumentów perkusyjnych, itd. Mikrofony mają wtedy specjalnie dobrane charakterystyki widmowe, tak aby najlepiej wydobyć w sygnału pożądane częstotliwości.

Czułość mikrofonu jest wyznacznikiem napięcia generowanego przez mikrofon przy danym ciśnieniu akustycznym. Mikrofony pojemnościowe mają czułość ok. 5mV/Pa, dynamiczne wstęgowe – 1mV/Pa, a dynamiczne cewkowe – 2mV/Pa.

Impedancja mikrofonu oznacza efektywną rezystancję wyjściową przy częstotliwości 1kHz. Przeważnie zawiera się w granicach 150-600Ω, w zależności od mikrofonu.



Rysunek 16.

Przykładowa charakterystyka widmowa mikrofonu

2.1.4 TECHNIKI MIKROFONOWE

Technikami mikrofonowymi nazywa się dobór rodzaju i sposobu ustawienia mikrofonów przy nagrywaniu, które umożliwiają uzyskanie pożądanego efektu. Obok zastosowania mikrofonu do każdego instrumentu z osobna, najczęściej stosuje się stereofoniczne techniki mikrofonowe, mające na celu odwzorowanie lokalizacji źródeł dźwięku.

Technika WM polega na zastosowaniu oddzielnego mikrofonu dla każdego źródła dźwięku. Dzięki ustawieniu mikrofonu blisko źródła, minimalizuje się wpływ pomieszczenia na sygnał. Ponadto, gdy sygnał pochodzi z wielu różnorodnych źródeł (np. orkiestra symfoniczna), takie ustawienie umożliwia ograniczenie wpływu poszczególnych rodzajów źródeł na siebie. Liczba oraz sposób ustawienia mikrofonów w tej technice zależy od instrumentu, który ma być rejestrowany.

Technika XY polega na ustawieniu na jednym statywie dwóch mikrofonów o charakterystyce kardoidalnej lub ósemkowej pod kątem +/- 45 stopni do osi frontalnej. Lokalizacja źródła jest zdeterminowana różnicą napięć w poszczególnych kanałach. Mikrofony powinny być tego samego typu i producenta. Kąt rozstawu osi mikrofonów powinien zawierać się w granicach 60-135° (najczęściej 90°). Użycie mikrofonów o charakterystyce ósemkowej umożliwia pełniejsze odtworzenie klimatu akustycznego panującego w danym pomieszczeniu.

Technika MS (ang. *Middle and Side*) – w tym przypadku używa się dwóch mikrofonów: jednego o charakterystyce kardoidalnej, drugiego o ósemkowej lub dwóch mikrofonów o charakterystyce ósemkowej. Mikrofon M (*middle*, kardioda lub ósemka) zbiera dźwięk bezpośredni, zaś S (*side*, ósemka) – dźwięk odbity oraz otaczający. Zmieniając proporcje między poziomami sygnałów z mikrofonów M i S można uzyskiwać różne efekty brzmieniowe, polegające na zmianie panoramy nagrania.

W **technice AB** ustawia się dwa mikrofony w odległości od 17 cm do 1,5 m między nimi. Mikrofony powinny mieć taką samą charakterystykę i być tego samego typu. W systemie tym wykorzystuje się różnice czasowe i amplitudowe sygnałów dochodzących do poszczególnych mikrofonów, aby wytworzyć obraz stereofoniczny. Wadą tej techniki jest wrażliwość na różnice faz w poszczególnych kanałach.

Zadanie 2. Nagrania z użyciem różnych typów mikrofonów

Do wykonania zadania jest potrzebny:

- komputer z zainstalowaną kartą dźwiękową i zestawem głośników komputerowych,
- oprogramowanie do obsługi karty dźwiękowej,
- mikrofon elektretowy, dynamiczny, pojemnościowy,
- wzmacniacz mikrofonowy i/lub prosty mikser kanałów.

Należy zapoznać się z możliwościami karty dźwiękowej zainstalowanej w komputerze. W szczególności należy przyjrzeć się temu, jakie sygnały mogą być doprowadzone do karty, a jakie wyprowadzone. Na tym etapie słuchacz powinien wiedzieć jakiego rodzaju urządzenia można do karty dźwiękowej dołączyć i czy można przez niewłaściwe połączenia coś zepsuć. Ponieważ w dalszej części będziemy używali mikrofonów należy zastanowić się gdzie je przyłączyć i w jaki sposób.



Następnie należy zapoznać się z możliwościami karty dźwiękowej w zakresie przetwarzania dźwięków. Większość efektów dźwiękowych jest możliwa do uzyskania i przetestowania przy użyciu standardowego oprogramowania dostarczonego z kartą. Słuchacz powinien wiedzieć jakie efekty dźwiękowe można uzyskać. Wszystkie efekty dźwiękowe należy przetestować.

W dalszej części przetestujemy standard MIDI. Większość programów do obsługi kart dźwiękowych ma możliwość uruchomienia klawiatury programowej. Korzystając z syntezy wavetable można zasymulować pianino i inne instrumenty, jak na przykład gitarę, perkusję. Należy użyć syntezy wavetable i, jeśli jest dostępna, syntezy FM i porównać jakość generowanego dźwięku. Do czego można wykorzystać standard MIDI?

Należy zapoznać się z wyglądem i budową mikrofonów: elektretowego, dynamicznego, pojemnościowego. Czy można poznać po wyglądzie rodzaj mikrofonu? Należy wiedzieć, który mikrofon można podłączyć bezpośrednio do karty dźwiękowej, a który wymaga zastosowania wzmacniacza mikrofonowego. Który typ mikrofonu wymaga dodatkowego zasilania?

Następnie należy zapoznać się z obsługą wzmacniacza mikrofonowego i miksera kanałów. W jaki sposób przyłączyć do nich mikrofony i jak dołączyć urządzenia do komputera? Na koniec należy wykonać wszystkie połączenia.

W następnym kroku spróbujemy zarejestrować sygnały dźwiękowe przy użyciu różnych mikrofonów i konfiguracji sprzętu. Przechwytywanie dźwięku z mikrofonów umożliwi program obsługi karty dźwiękowej. Można zrealizować dowolne nagrania. Przy okazji realizacji nagrań należy ustawić różne parametry nagrywania: rozdzielczość bitową i częstotliwość próbkowania. Jak te parametry wpływają na jakość nagranego dźwięku?

Na koniec należy porównać jakość nagrań dźwiękowych uzyskanych przy użyciu różnych typów mikrofonów. Które z tych nagrań jest najlepszej jakości i dlaczego?

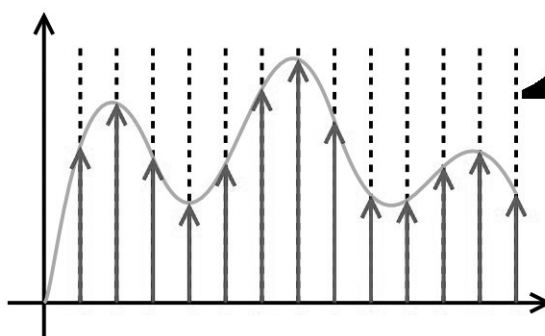
Jeśli wzmacniacz mikrofonowy / mikser kanałów umożliwia podłączenie dwóch mikrofonów monofonicznych jednocześnie należy zmiksować (połączyć) dwa kanały monofoniczne w jeden kanał. Czy w ten sposób możemy uzyskać dźwięk stereofoniczny i czy umiejscowienie mikrofonów ma w tym wypadku duże znaczenie?

2.2 FORMATY ZAPISU I PRZECHOWYWANIA PLIKÓW MULTIMEDIALNYCH

Pliki przechowujące materiały multimedialne często muszą umożliwić zapis i przechowywanie różnego rodzaju danych: dźwięku, obrazów, filmu, napisów, informacji o rozdziałach itp. Potrzebny jest do tego specjalny format zapisu danych, który będzie umożliwiał poprawne wyświetlenie lub synchronizację danych w celu ich jednoczesnego odtworzenia. Ogólnie taki format zapisu nazwa się **kontenerem multimedialnym**.

Dźwięk przechowywany w kontenerze multimedialnym musi być zapisany w jakiejś postaci cyfrowej. Jedną z najpopularniejszych metod zapisu sygnału dźwiękowego jest **PCM** (ang. *Pulse Code Modulation*). Metoda ta jest używana w telekomunikacji, w cyfrowej obróbce sygnału (np. w procesorach dźwięku), do zapisu na płytach CD (CD-Audio) i w wielu zastosowaniach przemysłowych.

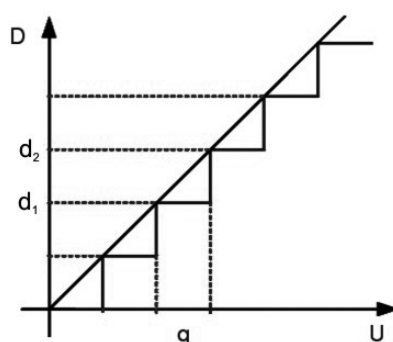
Metoda PCM polega na reprezentacji wartości chwilowej sygnału (próbkowaniu) w określonych (najczęściej równych) odstępach czasu (rys. 17), czyli z określoną częstością (tzw. częstotliwość próbkowania).



Rysunek 17.

Ilustracja zapisu dźwięku w formacie PCM

Wartość chwilowa sygnału jest przedstawiana za pomocą słowa kodowego, którego wartości odpowiada ją wybranym przedziałom kwantyzacji sygnału wejściowego. Przydział zakresu wartości analogowej jednej wartości cyfrowej jest nazywany kwantyzacją sygnału, prowadzi to do pewnej niedokładności (błąd kwantyzacji). Ilustracja kwantyzacji jest przedstawiona na rys. 18. Z konkretnego przedziału kwantyzacji q wartości analogowe z przedziału od d_1 do d_2 zostaną zastąpione jedną wartością zapisaną cyfrowo najbliższą liczbie d_1 . Liczba poziomów kwantyzacji jest zazwyczaj potęgą liczby 2 (ponieważ do zapisu próbek używane są słowa binarne) i wyraża się wzorem 2^n , gdzie n to liczba bitów przeznaczona na pojedynczą próbkę. Im większa częstotliwość próbkowania i im więcej bitów słowa kodowego reprezentuje każdą próbkę, tym dokładność reprezentacji jest większa, a tak zapisany sygnał jest wierniejszy oryginałowi. Dobór częstotliwości próbkowania w taki sposób aby połowa częstotliwości próbkowania (częstotliwość Nyquista) była większa od najwyższej częstotliwości składowej sinusoidalnej występującej w sygnale dźwiękowym (analiza widmowa), umożliwi bezstratną informacyjnie zamianę sygnału ciągłego na dyskretny.



Rysunek 18.
Kwantyzacja sygnału

Dźwięk w formacie PCM może być zapisywany z różną częstotliwością próbkowania, najczęściej jest to 8kHz (niektóre standardy telefonii), 44.1kHz (płyty CD-Audio) oraz różną rozdzielczością, najczęściej 8, 16, 20 lub 24 bity na próbkę, może reprezentować 1 kanał (dźwięk monofoniczny), 2 kanały (stereofonia dwukanałowa) lub więcej (stereofonia dookólna). Reprezentacja dźwięku próbkowana z częstotliwością 44.1kHz i w rozdzielczości 16 bitów na próbkę (65536 możliwych wartości amplitudy fali dźwiękowej na próbkę) jest uważana za bardzo wierną swemu oryginałowi, ponieważ z matematycznych wyliczeń wynika, iż pokrywa cały zakres pasma częstotliwości słyszalnych przez człowieka oraz prawie cały zakres rozpiętości dynamicznej słyszalnych dźwięków. Taki format kodowania zastosowano na płytach CD-Audio.

Inne formy cyfrowego kodowania dźwięku są zazwyczaj dużo bardziej złożone. Często wykorzystują różne metody kompresji danych w celu zredukowania ich ilości. Istnieją 2 rodzaje kompresji:

- **kompresja bezstratna** – algorytm upakowania informacji do postaci zawierającej mniejszą liczbę bitów w taki sposób, aby informację dało się odtworzyć do postaci identycznej z oryginałem,
- **kompresja stratna** – algorytm zmniejszania ilości bitów potrzebny do wyrażenia danej informacji, przy czym nie ma gwarancji, że odtworzona informacja będzie identyczna z oryginałem. Dla niektórych danych algorytm kompresji stratnej może odtworzyć informację prawie idealnie.

Przetworzenie pliku dźwiękowego do określonego formatu cyfrowego wymaga specjalnego programu, tzw. **kodeka**, w którym zaimplementowane są zaawansowane algorytmy cyfrowego przetwarzania sygnałów dźwiękowych. Poniżej krótko opisano najpopularniejsze kodeki dźwięku. W dalszej części szerzej będzie opisany sposób kodowania MP3.

Ogg Vorbis jest kodekiem ogólnego zastosowania. Najlepiej sprawdza się w tworzeniu plików o dużym stopniu kompresji (od 48 do 128kbps). Uznaje się, że średnia jakość dźwięku zakodowanego w formacie Ogg Vorbis jest porównywalna do AAC i wyższa niż MP3 o tej samej przepływności (czyli szybkości transmisji danych mierzonej w bitach na jednostkę czasu). W odróżnieniu od MP3 format Ogg Vorbis nie jest opatentowany i po-



zostaje bezpłatny, zarówno do celów prywatnych, jak i komercyjnych. Dekodowanie plików zapisanych w tym formacie wymaga większego zapotrzebowania na moc obliczeniową procesora niż MP3 (w przenośnych odtwarzaczach szczególnie uwidacznia się to poprzez skrócenie czasu pracy). Jest kodekiem z natury typu VBR (czyli dźwięk jest kodowany ze zmienną w czasie szybkością przepływu danych).

MPEG-4 Part 14 został utworzony w oparciu o format kontenera Apple QuickTime i jest właściwie identyczny z formatem MOV, ale wspiera wszystkie właściwości standardu MPEG. Pliki z zakodowanym dźwiękiem mają często rozszerzenie .mp4, nie istnieje natomiast coś takiego jak format kompresji dźwięku MP4.

AAC (ang. *Advanced Audio Coding*) to z kolei algorytm stratnej kompresji danych dźwiękowych, którego specyfikacja została opublikowana w roku 1997. Format AAC został zaprojektowany jako następcą MP3, oferujący lepszą jakość dźwięku przy podobnym rozmiarze danych. Kompresja AAC jest modułarna i oferuje w standardowo cztery profile:

- **Low Complexity (LC)** – najprostszy, najszerzej stosowany i odtwarzany przez wszystkie odtwarzacze obsługujące format AAC,
- **Main Profile (MAIN)** – rozszerzenie LC,
- **Sample-Rate Scalable (SRS)** lub **Scalable Sample Rate (AAC-SSR)** – zakres częstotliwości dzielony jest na cztery kompresowane niezależnie pasma, jakość jest przez to nieco niższa niż pozostałych profili,
- **Long Term Prediction (LTP)** – rozszerzenie MAIN wymagające mniejszej ilości obliczeń.
- Usprawnienia AAC w stosunku do poprzednich algorytmów kompresji dźwięków
- próbkowanie 8-96kHz (MP3 16-48kHz),
- do 48 kanałów (MP3 – 2 kanały w standardzie MPEG-1 i 5.1 w standardzie MPEG-2),
- skuteczniejszy i wydajniejszy,
- lepsze przenoszenie częstotliwości ponad 16kHz ,
- lepszy tryb kompresji sygnału stereofonicznego joint-stereo.

2.3 CYFROWE PRZETWARZANIE SYGNAŁÓW DŹWIĘKOWYCH

Większość efektów cyfrowego przetwarzania sygnałów dźwiękowych przydatnych w domowym studiu nagrań można uzyskać przy użyciu dobrej karty dźwiękowej. Lepszą jakość, większe możliwości zarówno przetwarzania jak i montażu dźwięku w zastosowaniach profesjonalnych i amatorskich oferują jednak urządzenia zewnętrzne.

Podstawowym urządzeniem jest **stół mikserski**. Jego zadaniem jest dopasowanie poziomów sygnałów fonicznych pochodzących z różnych źródeł. Do podstawowych operacji, jakie można na sygnale wykonać, należą: regulacja poziomu, miksowanie oraz przetaczanie. Z tych wymienionych tylko regulacja poziomu jest ściśle przyporządkowana pojedynczemu torowi obróbki. Dwie pozostałe operacje dokonują przemieszania pomiędzy torami. W każdym stole występuje także część operacyjna z elementami wykonującymi bardziej skomplikowane czynności.

Następna grupę stanowią procesory dźwięku. Można tu wyróżnić następujące urządzenia:

- **procesory dynamiki** – urządzenia do kształtowania zakresu dynamiki sygnału w celu dostosowania go do możliwości toru fonicznego lub poprawienia akustycznych parametrów sygnału wśród nich można wyróżnić:
 - **kompresor** – umożliwia zmniejszenie dynamiki sygnału poprzez zmniejszenie wzmocnienia głośniejszych sygnałów przy pozostawieniu oryginalnego wzmocnienia sygnałów cichych,
 - **limiter** – działa podobnie jak kompresor z tym, że nie przepuszcza praktycznie sygnałów o poziomie większym niż pewien próg,
 - **ekspander** – działa odwrotnie do kompresora (w zasadzie nie jest wykorzystywany oddzielnie, tylko w powiązaniu z kompresorem tworzą układ zwany kompanderem);
- **bramka szumowa** – służy do usuwania niepożądanego szumu; jeżeli sygnał spada poniżej określonego poziomu, nie pojawia się na wyjściu;
- **układ automatycznej regulacji wzmocnienia** – zmienia poziom sygnału do pożądanej wartości (tzn. wzmacnia sygnały o poziomie niższym od pożądanej wartości, a komprimuje te o poziomie wyższym);
- **procesory efektowe** – służą do zmiany właściwości czasowych i przestrzennych sygnału, a także zmiany wysokości dźwięku np.:
 - **pogłos (reverb)** i **opóźnienie (delay)** – kształtują właściwości czasowe sygnału. Pierwszy z nich polega na dodaniu do sygnału wygenerowanych losowych ech, które następują po sobie na tyle



szybko i jest ich na tyle dużo, że ucho ludzkie traktuje całość jako stopniowo zanikający sygnał. Opóźnienie polega na dodaniu sygnału wejściowego opóźnionego o wybraną wartość czasową oraz z wybranym poziomem do nieopóźnionego sygnału wejściowego,

- **zmiana wysokości dźwięku (*pitch-shift*)** – polega na przepróbkowaniu sygnału wejściowego (zwiększenie lub zmniejszenie częstotliwości próbkowania), a następnie transmisji tak przetworzonego sygnału z ustaloną częstotliwością próbkowania;
- **wzmacniacze zewnętrzne** – oprócz przedwzmacniacza w stole mikserskim, stosuje się również: przedwzmacniacze do mikrofonów – wyrównują poziom sygnału z mikrofonu do poziomu sygnałów liniowych, wzmacniacze mocy – umożliwiają odpowiednie wzmocnienie sygnału podawanego do monitorów głośnikowych.

Sygnały po ewentualnej wstępnej korekcji są następnie rejestrowane za pomocą jednego z urządzeń zapisujących. Po zakończeniu faz rejestracji i dogrywania rozpoczyna się faza miksowania. Podczas miksowania uprzednio zapisane ścieżki w rejestratorze wielościeżkowym są wybierane i łączone przy użyciu stołu mikserskiego. Po zmiksowaniu w dół wszystkich utworów projektowanego nagrania (np. do zapisu stereo), materiał muzyczny jest poddawany dalszej obróbce, czyli edycji, masteringowi i powielaniu.

Zadanie 3. Przetwarzanie i montaż sygnału dźwiękowego

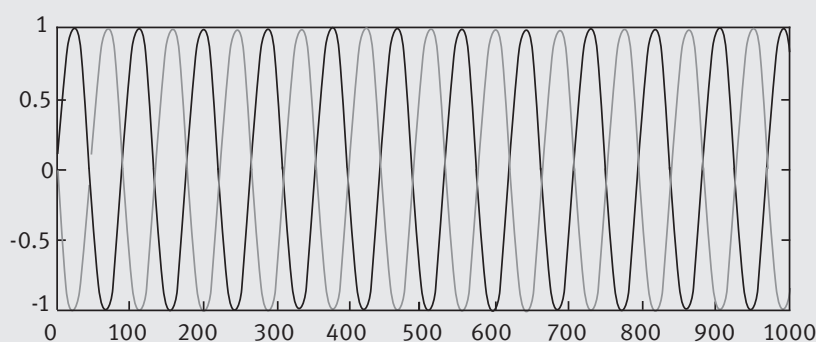
Do wykonania zadania jest potrzebny:

- komputer z zainstalowaną kartą dźwiękową i zestawem głośników komputerowych,
- program do przetwarzania dźwięku Audacity.

W tym ćwiczeniu poznamy, jak powstają proste efekty dźwiękowe, co można uzyskać przez zastosowanie filtrów cyfrowych w obróbce dźwięku i jak przeprowadzić prosty montaż sygnału dźwiękowego. Program Audacity to zaawansowany i wielościeżkowy edytor plików dźwiękowych, a co najważniejsze dostępny jest za darmo. Program ma rozbudowane menu efektów – dostępny jest między innymi kompresor, echo, podbicie basów, equalizer, filtry, odszumiacz, wyciszanie i wiele innych. Na początku należy zapoznać się z obsługą tego programu i jego możliwościami.

W następnym kroku zapoznamy się w techniką tworzenia prostych efektów dźwiękowych. Efekty te wykorzystują łączenie dwóch lub więcej kanałów sygnału dźwiękowego.

Najprostszym efektem dźwiękowym, który będzie zademonstrowany przez prowadzącego, jest dodanie do siebie dwóch sygnałów sinusoidalnych będących w przeciwfazie (rys. 19). W momencie kiedy dodamy do siebie sygnały elektryczne otrzymamy sygnał wynikowy zero. Co uzyskamy jeśli dodamy do siebie sygnały akustyczne? Jeśli kanał lewy odtwarza jedną sinusoidę, a kanał prawy drugą (przesuniętą o 180°) to po zbliżeniu głośniczków do siebie można usłyszeć, że głośność dźwięku wyraźnie spada. Sygnały akustyczne również podlegają sumowaniu. Nie uzyskamy jednak w ten sposób ciszy.

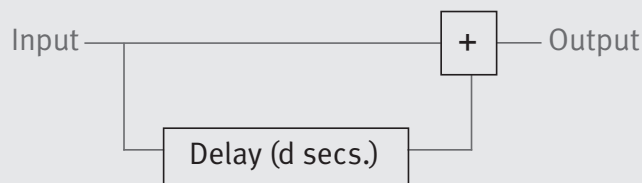


Rysunek 19.

Dwa sygnały sinusoidalne będące w przeciwfazie

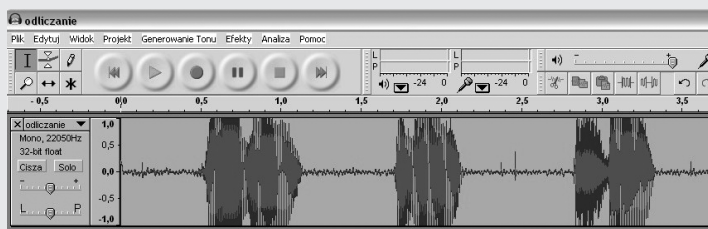
W następnym kroku słuchacze samodzielnie stworzą efekt echa i pogłosu. Jeśli chcemy stworzyć efekt echa, to do sygnału oryginalnego należy dodać sygnał opóźniony o mniejszej amplitudzie (sygnał podawany jest jeden raz). Ilustracja sposobu postępowania podana na rys. 20.





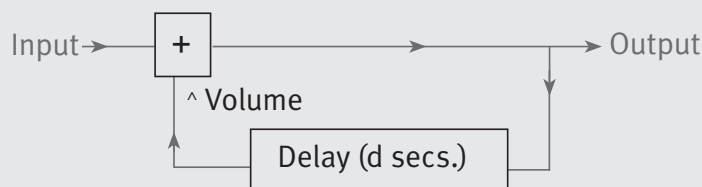
Rysunek 20.
Ilustracja graficzna tworzenia efektu echa

W celu stworzenia efektu echa w programie Audacity należy wgrać nagrany wcześniej sygnał odliczania do pięciu do programu (Plik ® Otwórz). Na ekranie zostanie wyświetlony przebieg czasowy tego sygnału (rys.21). Po lewej stronie wyświetlone są parametry z jakimi nagrany był dźwięk. Używając myszki można zaznaczać dowolne fragmenty nagrania. Z lewej strony z góry widoczna jest paleta z narzędziami. Wykorzystując narzędzia z palety można zaznaczać, przesuwać, wycinać fragmenty sygnałów. Narzędzie obwiednia służy do zmiany głośności sygnału. Żeby uzyskać efekt echa należy zaznaczyć cały, wgrany wcześniej, sygnał i skopiować go do nowej ścieżki audio (Projekt ® Nowa ścieżka audio). Następnie korzystając z narzędzia przesuwania w czasie, należy przesunąć cały sygnał w nowej ścieżce w stosunku do sygnału pierwotnego. Narzędzie obwiednia pozwoli dostosować głośność sygnału. Po odtworzeniu obu ścieżek na raz uzyskujemy efekt echa. Efekt echa będzie brzmiał bardzo różnie w zależności od stopnia przesunięcia i głośności sygnału w stworzonym nowym kanale. Gotowy efekt należy zapisać w pliku na dysku.



Rysunek 21.
Okno programu Audacity z wgranym sygnałem dźwiękowym

Pogłos uzyskuje się w podobny sposób. W tym przypadku sygnał opóźniony dodawany jest w pętli wiele razy, za każdym razem sumuje się on z poprzednim sygnałem wyjściowym i ma coraz mniejszą amplitudę (rys. 22). W programie Audacity możemy uzyskać efekt pogłosu dodając wiele kanałów dźwiękowych, w których sygnały dźwiękowe będą miały różne, losowe, przesunięcia (np. na początku można dodać pięć nowych kanałów). Amplitudy sygnałów dodawanych powinny być mniejsze od amplitudy sygnału wejściowego. Odpowiednio dobierając przesunięcia i amplitudy sygnałów dodawanych można na przykład uzyskać akustykę sali koncertowej lub małego pokoju. Oczywiście bez skomplikowanych obliczeń jest to bardzo trudne do osiągnięcia. Słuchacze mogą spróbować zmieniając przesunięcia i głośności sygnałów dodawanych uzyskać jakiś określony efekt (w tym przypadku, najprawdopodobniej będzie to jakiś efekt losowy).



Rysunek 22.
Ilustracja graficzna tworzenia efektu pogłosu

Następnie zapoznamy się z działaniem różnych wbudowanych w program Audacity filtrów i efektów dźwiękowych. W programie Audacity dostępnych jest bardzo wiele filtrów i efektów dźwiękowych. Uruchamia się je poleceniem Efekty. Dostępny jest między innymi filtr do tworzenia echa. Należy sprawdzić czy echo będące wynikiem działania tego filtru brzmi podobnie do stworzonego wcześniej „ręcznie” efektu. Słuchacze powinni wypróbować dostępne filtry i efekty. Należy zaobserwować, co się zmienia w brzmieniu dźwięku po zastosowaniu filtru i do czego ewentualnie uzyskany efekt brzmieniowy można użyć. W następnym kroku będziemy wykorzystywać niektóre efekty podczas montażu dźwięku.

W ostatnim etapie ćwiczenia spróbujemy zmontować nagranie dźwiękowe. Może to być własna, bardzo krótka (kilkuminutowa) audycja. Do nagrania wykorzystamy wcześniej zarejestrowane z użyciem mikrofonów pliki dźwiękowe (można je też nagrać teraz) oraz pliki dostarczone przez prowadzącego. Pliki te należy wcześniej odsłuchać. Zawierają one różne odgłosy i tła dźwiękowe. Na tym etapie słuchacze powinni stworzyć sobie jakąś koncepcję nagrania.

W celu przećwiczenia niektórych operacji pomocnych przy montażu dźwięku najpierw wykonamy kilka prostych ćwiczeń. Pierwsze z nich dotyczy przestawiania fragmentów sygnału dźwiękowego. Przestawianie wykonamy na pliku dźwiękowym z nagraniem odliczaniem. Należy tak poprzestawiać fragmenty sygnału dźwiękowego, żeby odliczanie odbywało się wstecz (5 4 3 2 1 0). Przestawianie realizujemy przez zwykłą operację wytnij/wstaw. Następne ćwiczenie dotyczy dostosowanie poziomów głośności sygnałów dźwiękowych. W nagraniu finalnym poziomy głośności poszczególnych elementów nagrania muszą ze sobą współgrać (czyli przede wszystkim być na podobnym poziomie), należy je więc odpowiednio dostosować. Dostosowanie poziomu głośności sygnałów można np. uzyskać przez wybranie polecenia Efekty ® Wzmacniaj. Nakładanie na siebie różnych ścieżek dźwiękowych wykonywaliśmy już to przy okazji tworzenia efektu echa i pogłosu. W przypadku montażu mogą być pomocne różne efekty wyciszania.

Po wykonaniu powyższych ćwiczeń można przejść do wykonania własnego nagrania. Efekt finalny należy zapisać w pliku .wav. Przykład krótkiego zmontowanego nagrania będzie zaprezentowany przez prowadzącego.



2.4 PSYCHOAKUSTYKA – PODSTAWOWE ZASADY KOMPRESJI SYGNAŁÓW DŹWIĘKOWYCH

Psychoakustyka to współczesna dziedzina wiedzy zajmująca się związkiem obiektywnych (fizycznych) cech dźwięku z jego cechami subiektywnymi, z wrażeniem jakie w mózgu słuchacza wywołują bodźce dźwiękowe. Psychoakustyka próbuje przewidzieć zachowanie się słuchu człowieka w określonych warunkach fizycznych.

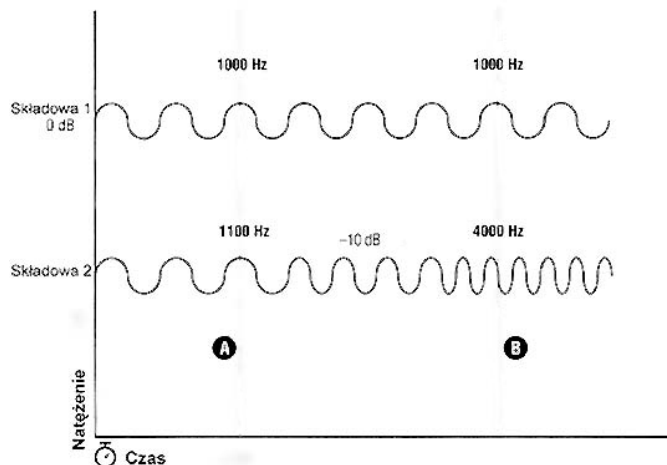
Modelami psychoakustycznymi nazywamy modele systemu słyszenia, które uwzględniają ograniczenia i tolerancje mechanizmów percepcji przeciętnego słuchacza, są to modele matematyczne mówiące jakie dźwięki są rozpoznawalne przez ludzkie ucho, jakie natomiast nie są. Modele psychoakustyczne są podstawą między innymi kompresji dźwięku, algorytmów oceny jakości transmisji mowy, systemów automatycznie rozpoznających mowę oraz systemów rozpoznających mówców.

Wytyczne do modelowania pochodzą z pomiarów psychoakustycznych (odsłuchowych), w których słuchacze oceniają wrażenia wywołane różnymi sygnałami testowymi prezentowanymi w określonym kontekście (np. czy słyszą ton sinusoidalny prezentowany na tle szumu). Model przetwarza sygnał w taki sposób, aby jego wyjście stanowiło predykcję subiektywnych ocen słuchaczy. Najprostszym faktem psychoakustycznym jest różna czułość ludzkiego ucha na dźwięki o różnych częstotliwościach (niektórych częstotliwości np. bardzo wysokich lub bardzo niskich nie słyszymy w ogóle). Modele psychoakustyczne przewidują zwykle zakres słyszalności od 20Hz-20kHz (dlatego właśnie większość współczesnych odtwarzaczy muzyki zapisanej cyfrowo ma takie pasmo przenoszenia) i maksymalną czułość w zakresie od 2kHz do 4kHz.

Innym szeroko stosowanym faktem psychoakustycznym jest **maskowanie dźwięków**. Najogólniej maskowanie polega na przysłanianiu sygnałów słabszych sąsiadujących z sygnałami znacznie głośniejszymi, które je zagłuszają. Rozróżniamy 2 rodzaje maskowania:

- maskowanie równoczesne,
- maskowanie czasowe.

Efekt maskowania równoczesnego opiera się na fakcie, że człowiek nie jest w stanie odróżnić dwóch dźwięków o zbliżonej częstotliwości, jeśli jeden z nich jest znacznie głośniejszy od drugiego (rys. 23 przypadek A). Możliwe jest to dopiero wtedy, gdy sygnały mają zupełnie różne częstotliwości (przypadek B).



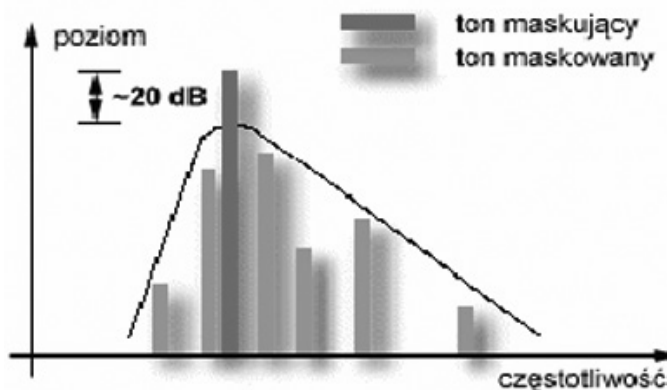
Rysunek 23.

Efekt maskowania równoczesnego

Najprościej mówiąc maskowanie równoczesne polega na tym, że ciche dźwięki o częstotliwościach zbliżonych do częstotliwości dźwięku głośniego nie są słyszalne. Wszystkie standardy MPEG audio (a więc również MP3) wykorzystują tę właściwość ucha ludzkiego, bazując one na usuwaniu słabszych dźwięków, które nie docierają do mózgu człowieka.

Maskowanie czasowe polega na eliminacji składowych o mniejszym natężeniu, które mają zbliżoną częstotliwość do dźwięku o większym natężeniu i występują razem w pewnym przedziale czasu. Na rys. 24 jest pokazany efekt maskowania czasowego – czarną linią zaznaczono próg słyszalności. Można w tym przypadku wyróżnić dwa typy maskowania:

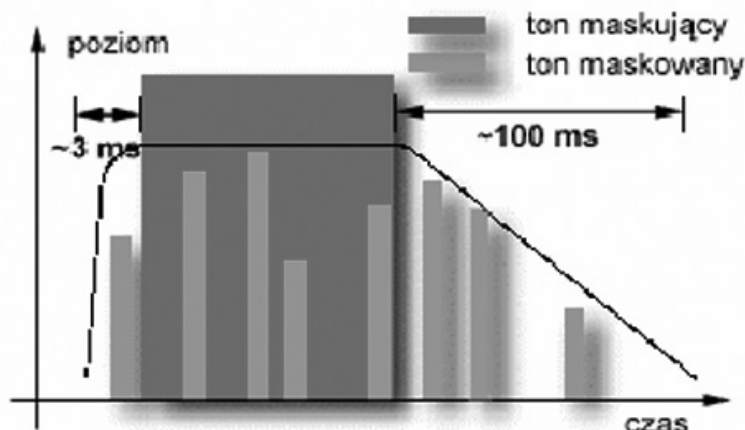
- maskowanie dźwięków następujących (maskowanie pobudzcowe) – głośny dźwięk potrafi zagłuszyć cichsze dźwięki następujące zaraz po nim,
- maskowanie dźwięków poprzedzających (maskowanie wsteczne) – cichy dźwięk poprzedzający w krótkim czasie dźwięk głośniejszy nie jest słyszalny. Ta własność układu słuchowego jest szczególnie ciekawa, gdyż nie da się jej wyjaśnić na gruncie adaptacji krótkoterminowej układu słuchowego. Równocześnie pokazuje ona, że układ słuchowy nosi pewne cechy układu nieprzyczynowego (tzn. skutek wywołany przez jakiś bodziec występuje przed wystąpieniem bodźca).



Rysunek 24.

Efekt maskowania czasowego

Na rys. 25 zilustrowano efekt maskowania równoczesnego i czasowego jednocześnie. Czarna linia oznacza próg słyszalności. Słabe dźwięki (kolor zielony), które są maskowane przez dźwięk silniejszy, mogą zostać podczas kompresji usunięte. Pozostanie tylko dźwięk słyszalny (kolor czerwony).



Rysunek 25.

Całkowity efekt maskowania

2.5 IDEA KOMPRESJI MP3

W 1987r. w niemieckim instytucie Fraunhofer rozpoczęto prace nad radiofonią cyfrową. Jednym z kluczowych elementów było opracowanie systemu kompresji danych umożliwiającego skuteczny zapis sygnałów dźwiękowych. Algorytmy opracowane w instytucie Fraunhofer stały się później podstawą systemu MP3.

Należy zaznaczyć, że algorytm stosowany przy kompresji MP3 wykorzystuje kompresję stratną – przy odtwarzaniu, dźwięk nie odpowiada dokładnie dźwiękowi sprzed kompresji. Kompresja powoduje nawet ponad dziesięciokrotne zmniejszenie wielkości miejsca na dysku w stosunku do objętości dźwięku, który kompresji nie podlegał. Jednak osoby z bardziej wrażliwym słuchem odbierają dźwięk skompresowany jako gorszy pod względem jakości.

W roku 1991 prace nad algorytmem kodowania MPEG-1 – Layer3 opracowywanym w instytucie Fraunhofer zostały ukończone. Opracowany algorytm stał się najbardziej optymalnym sposobem kodowania sygnałów audio w rodzinie określonej przez międzynarodowe normy ISO-MPEG. Używając tego algorytmu (znanego powszechnie w Internecie jako MP3, ze względu na rozszerzenie) do kodowania plików audio, jakość „prawie CD”, tj. stereo, 44KHz, 16 bitów, można uzyskać przy przepływności 112 – 128kbps (stopień kompresji 11:1 – 13:1).

Kompresja MP3 jest oparta na matematycznym modelu psychoakustycznym ludzkiego ucha.

- Idea kompresji MP3 polega na wyeliminowaniu z sygnału tych danych, które są dla człowieka niesłyszalne lub które słyszymy bardzo słabo.
- Kompresja MP3 jest połączeniem metody kompresji stratnej z kompresją bezstratną.
- Etap 1 – koder eliminuje z sygnału składowe słabo słyszalne i niesłyszalne dla człowieka (kompresja stratna).
- Etap 2 – uzyskane dane są poddawane dodatkowej kompresji w celu eliminacji nadmiarowości (kompresja bezstratna).

Algorytm operuje na dźwięku próbkowanym z jakością: 16; 22.5; 24; 32; 44.1 oraz 48kHz. Jest optymalizowany pod wyjściową przepustowość 128kbps dla sygnału stereo, aczkolwiek dostępne są przepustowości od 32kbps do 320kbps.



Algorytm kodowania MP3 może operować na 4 rodzajach dźwięku wejściowego:

- *mono*,
- *stereo* – kompresja dwóch oddzielnych strumieni,
- *joint stereo* – badane jest podobieństwo sygnałów w obu kanałach, jeśli w obu kanałach jest ten sam sygnał, to koder przełącza się do trybu mono, umożliwia to kodowanie dźwięku z większą dokładnością,
- *dual channel* – zawiera dwa niezależne kanały, jest stosowany np. przy tworzeniu kilku różnych wersji językowych dla filmu.

W procesie kodowania MP3 występuje kilka procesów, które wymagają dodatkowego wyjaśnienia. Należą do nich dyskretna transformacja kosinusowa, kwantyzacja, kodowanie Huffmana.

Dyskretna transformacja kosinusowa (DCT) pomaga rozdzielić sygnał na części, przekształcając dane do postaci umożliwiającej zastosowanie efektywnych metod kompresji. DCT przetwarza sygnał określony w dziedzinie czasu na sygnał określony w dziedzinie częstotliwości. W wyniku działania transformaty na sygnale wejściowym powstają odpowiadające mu współczynniki transformaty. Transformata kosinusowa jest odwracalna, to znaczy, że dysponując tylko współczynnikami transformaty można odtworzyć odpowiadający im sygnał bez żadnych strat. Zaletą transformaty DCT jest to, że większość współczynników jest zwykle bliska zeru, a zatem po procesie kwantyzacji współczynniki te można pominąć, co umożliwia lepszą kompresję danych.

Kwantyzacja jest to proces ograniczenia zbioru wartości sygnału w taki sposób, aby można go było zapisać na skończonej liczbie bitów. Polega na przypisaniu wartości analogowych do najbliższych poziomów reprezentacji, co oznacza nieodwracalną utratę informacji (rys. 6). Kwantyzacja polega na przeskalowaniu współczynników DCT poprzez podzielenie ich przez właściwy współczynnik znajdujący się w tabeli kwantyzacji, a następnie zaokrągleniu wyniku do najbliższej liczby całkowitej. Tablice kwantyzacji dobierane są doświadczeniowo.

Kodowanie Huffmana jest to bezstratna metoda kodowania, przedstawiona przez Davida Huffmana w roku 1952. Kodowanie Hoffmana stanowi jedną z najprostszych i łatwych w implementacji metod kompresji bezstratnej. W algorytmie jest wykorzystywany fakt, że pewne wartości danych występują częściej niż inne. Jeżeli zatem zakodujemy częściej występujące wielkości za pomocą krótszych słów kodowych, a rzadziej występujące – za pomocą dłuższych, to sumarycznie długość zakodowanych danych będzie krótsza niż przed kodowaniem.

2.6 KODOWANIE DŹWIĘKU W STANDARDZIE MP3

Poniżej podane są najważniejsze etapy kodowania dźwięku w standardzie MP3.

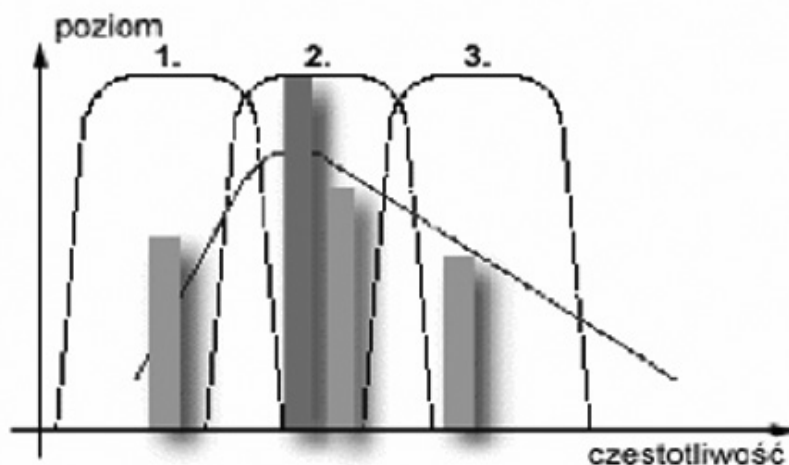
- Sygnał wejściowy jest dzielony na mniejsze fragmenty zwane **ramkami** o czasie trwania ułamka sekundy.
- Dla sygnału dźwiękowego określonego w czasie wyliczana jest jego reprezentacja częstotliwościowa, czyli wyliczane jest widmo sygnału dźwiękowego.
- Widmo sygnału dla każdej ramki porównywane jest z matematycznym modelem psychoakustycznym. W wyniku tego porównania koder określa, które ze składowych dźwięku jako najlepiej słyszalne muszą zostać odwzorowane najwierniej, a które można zakodować w przybliżeniu lub w ogóle pominąć.
- Ustalany jest optymalny przydział bitów na poszczególne częstotliwości pasma akustycznego, tak aby zapewnić możliwie najwierniejsze zakodowanie sygnału.
- Strumień bitów jest poddawany ponownej kompresji poprzez kodowanie Huffmana. Celem tej operacji jest usunięcie nadmiarowości z danych przetworzonych w pierwszym etapie, czyli dodatkowa kompresja bezstratna.
- Kolejne ramki poprzedzone nagłówkami są składane w pojedynczy ciąg bitów (strumień bitowy). Nagłówki zawierają metainformacje określające parametry poszczególnych ramek.

Kompresja MP3 rozpoczyna się rozdzieleniem sygnału wejściowego na małe fragmenty trwające ułamek sekundy (zwane inaczej ramkami), oraz podział tych fragmentów według pasma na 576 części – najpierw 32 w wielofazowym banku filtrów, a następnie podpasma przekształcane są dyskretną transformatą kosinusową, która generuje 18 współczynników dla każdego podpasma. Zwiększa to szanse na usunięcie niepotrzebnych informacji, sygnał może też być lepiej kontrolowany w celu śledzenia progów maskowania.

Na rys. 26 zobrazowano ideę działania banku filtrów. Czarne linie oznaczają podział sygnału dźwiękowego na pasma częstotliwościowe 1, 2 i 3. Niebieska linia wyznacza poziom progów słyszalności wyliczony



na podstawie modelu psuchoakustycznego. Dwa “zielone” sygnały po prawej stronie znajdują się poniżej poziomu słyszalności. Można więc usunąć sygnał w trzecim podzakresie. Sygnał z lewej strony jest słyszalny, można jednak podnieść dopuszczalny poziom szumów, czyli zapisać go mniejszą liczbą bitów. Jeśli kwantowany dźwięk da się utrzymać poniżej progu maskowania, to efekt kompresji powinien być nieodróżnialny od oryginalnego sygnału.



Rysunek 26.

Idea działania banku filtrów

Proces kwantyzacji w kompresji MP3 jest realizowany na zasadzie dwóch pętli, jedna zagnieżdżona w drugiej (rys. 27). Zawiera on także część procesu formowania dźwięku.

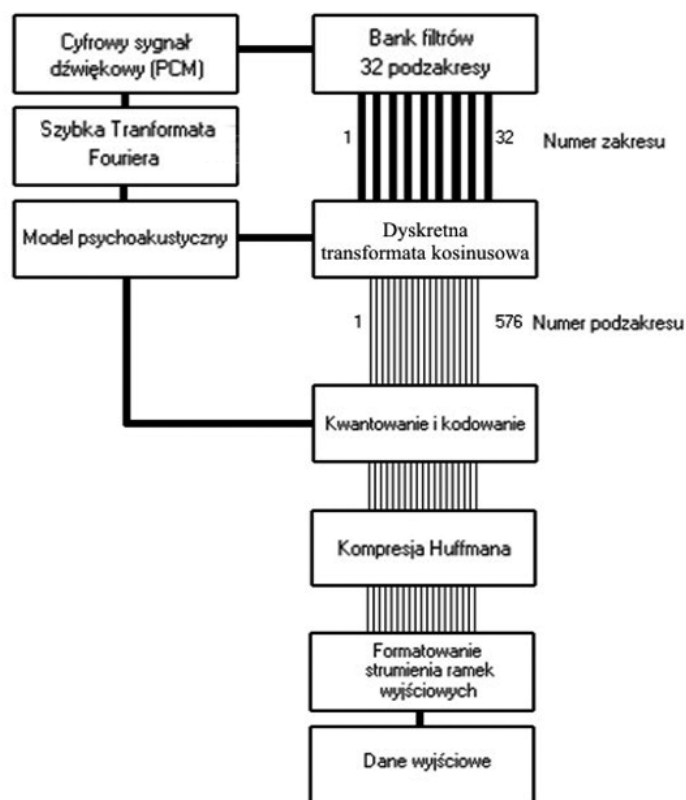
Pierwsza z pętli, wewnętrzna, to pętla kontroli współczynnika kompresji. Przeprowadzany jest w niej proces kwantyzacji dla poszczególnych pasm częstotliwościowych, następnie symulowane jest kodowanie skwantowanych współczynników. Jeżeli po kodowaniu okaże się, że jest przekroczony limit przepływności, czyli plik po kompresji byłby zbyt duży, to wskaźnik przyrostu jest dopasowywany do danych i cała pętla jest powtarzana od nowa.

Druga pętla, zewnętrzna, pętla kontroli zniekształceń rozpoczyna się od ustawienia indywidualnych współczynników kwantyzacji na 1, po czym obliczany jest błąd kwantyzacji. Jeśli błąd ten przekracza oszacowany przez model psuchoakustyczny próg percepcji, to jest odpowiednio zmieniany współczynnik kwantyzacji i obliczenie błędu odbywa się ponownie. Gdy nie jest możliwe uzyskanie żądanej przepływności i spełnienie wymagań modelu psuchoakustycznego, to dźwięk jest kodowany mimo niespełnienia wymagań.

Po procesie kwantyzacji następuje proces kompresji algorytmem Huffmana. W celu dopasowania procesu kompresji do fragmentu danych źródłowych wybierana jest najbardziej pasująca tablica kodów Huffmana z całego zestawu. W celu otrzymania lepszego dopasowania, różne tablice kodów Huffmana są wybierane dla różnych części widma. Jest to proces usuwania nadmiarowych danych bez utraty informacji. Bazuje ono na słowie kodowym – kluczu o zmiennej długości, w której klucze krótkie przypisane są do często występujących wzorców, a długie do rzadko występujących. Algorytm rozpoczyna działanie od stworzenia histogramu (tablicy częstości występowania danych w pliku). W drugim kroku tworzy listę drzew binarnych, które w węzłach przechowują symbol i częstość jego wystąpienia. Następnie w pętli, dopóki jest jeszcze więcej niż jedno drzewo na liście usuwamy 2 drzewa, które mają w korzeniu zapisane najmniejsze zsumowane częstości i wstawiamy nowe drzewo, którego korzeń zawiera sumę częstości usuniętych drzew.

Końcowym etapem procesu kompresji jest formatowanie ramek wyjściowych i zapis do strumienia wyjściowego. Niektóre pliki MP3 dodatkowo zawierają sumy kontrolne. Suma kontrolna to 16 bitowa liczba, która jest zapisywana w każdej ramce oddzielnie i służy do weryfikacji poprawności strumienia MP3.





Rysunek 27.
Kodowanie MP3

STRUMIEŃ BITOWY

Gęstość strumienia bitowego (ang. *bitrate*) określa współczynnik kompresji sygnału algorytmem MP3. Wyznacza on liczbę bitów przypadającą na sekundę finalnego zapisu. Ustawienie odpowiedniej wartości strumienia bitowego jest kompromisem między jakością, a rozmiarem pliku wynikowego (rys. 28).



Rysunek 28.
Ilustracja pojęcia strumienia bitowego

Kompresja MP3 może przebiegać:

- ze stałą gęstością strumienia bitowego (ang. *constant bitrate*) – tryb CBR,
- zmienną gęstością strumienia bitowego (ang. *variable bitrate*) – tryb VBR.

tryb CBR – każda sekunda dźwięku skompresowana jest za pomocą tej samej liczby bitów, co powoduje jednak, że różne fragmenty utworu mają niejednakową jakość (spokojny fragment wykonany na instrument solo brzmi lepiej, niż „mocne uderzenie” całej orkiestry wspomaganą chórem),

tryb VBR – koduje sygnał uwzględniając jego dynamikę, dzięki czemu przydziela więcej bitów fragmentom sygnału, który zawiera dużo ważnych informacji, oraz mniej bitów dla części sygnału, które są mniej złożone. Każda sekunda dźwięku skompresowana jest za pomocą odpowiednio dobranej liczby bitów, dzięki czemu cały utwór ma stałą jakość. W tym wypadku spokojny fragment wykonany na instrument solo (dając się mocniej skompresować) brzmi tak samo dobrze, co „mocne uderzenie” całej orkiestry wspomaganą

chórem (wymagające mniejszego stopnia kompresji). Kompresja w trybie VBR wymaga podania przedziału tolerancji, w jakim może się zmieniać gęstość strumienia bitowego.

Ponieważ zadana gęstość strumienia bitowego obowiązuje dla każdej ramki, w przypadku bardzo złożonych fragmentów może okazać się niewystarczająca i program kodujący nie będzie w stanie zapewnić żądanej jakości zapisu w ramach przydzielonej liczby bitów. Aby zapobiec temu zjawisku standard MP3 zapewnia możliwość skorzystania z dodatkowej rezerwy umożliwiającej zapisanie nadmiarowych danych, tzw. **rezerwy bitowej**. Rezerwa ta powstaje w miejscu pustych fragmentów ramek, w których po zakodowaniu sygnału zostało trochę miejsca.

2.7 ŁĄCZENIE KANAŁÓW ZAPISU STEREOFONICZNEGO

Jak wiemy sygnał stereo składa się z dwóch odseparowanych od siebie kanałów. Przez znaczną część czasu kanały te jednak przenoszą jeśli nie identyczne to bardzo zbliżone do siebie informacje. Jeśli tak jest, to wtedy koder MP3 wykorzystuje tzw. algorytm joint-stereo, który powtarzające się dźwięki w obu kanałach zapisuje jako jeden.

Dodatkową możliwością podczas kodowania sygnału z funkcją joint stereo jest **stereofonia różnicowa**. Polega ona na zapisaniu dwóch ścieżek – kanału środkowego będącego sumą sygnałów R i L oraz kanał boczny będący ich różnicą, który służy później do rekonstrukcji sygnału oryginalnego podczas odtwarzania pliku. Warto dodać, że algorytm joint-stereo jest bardzo efektywny – umożliwia zredukowanie do 50 % ilości potrzebnych danych.

Ogólnie algorytm MP3 umożliwia skompresowanie dźwięku do postaci:

- **dual channel** – kanały lewy i prawy są traktowane jako dwa niezależne kanały mono, każdy z nich otrzymuje dokładnie połowę dostępnej przepływności; w praktyce nieekonomiczny więc nieużywany;
- **stereo** – kanały lewy i prawy są traktowane jako stereo, przepływność dzielona jest pomiędzy kanały dynamicznie (np. jeżeli w lewym kanale akurat jest cisza, to prawy dostaje większą część dostępnej przepływności – daje to lepszą jakość dźwięku w prawym kanale) – używany do kompresji w wysokich przepływnościach (192kbps i więcej);
- **joint stereo (stereofonia różnicowa)** – kanały lewy i prawy są rozbijane na kanały mid/side (mid = środek – to co jest identyczne w obu kanałach i side = otoczenie – to czym różnią się oba kanały) – używany do kompresji w średnich przepływnościach (128 – 192kbps);
- **intensity stereo** – kanały lewy i prawy są zamieniane na jeden kanał mono, do którego jest dodawana informacja o uśrednionym kierunku, z którego dźwięk dochodzi (dzięki czemu podczas odsłuchu dźwięk nie dochodzi ze środka tylko z jakiegoś kierunku) – używany do kompresji w niskich przepływnościach (128kbps i mniej);
- **mono** – kanały lewy i prawy są zamieniane na jeden kanał mono, który jest potem kompresowany, dźwięk odtwarzany jest jako mono – używany do bardzo niskich przepływności (32kbps i mniej), głównie do kompresji głosu.

Ciekawostką jest to, że specyfikacja formatu MP3 zawarta w dokumencie ISO/IEC 11172-3, nie określa dokładnie sposobu samego kodowania, a jedynie prezentuje ogólny zarys techniki, i określa wymagany poziom zgodności zapisu z normą. Innymi słowy, ustala ona kryteria, jakie musi spełniać struktura pliku, by można było go sklasyfikować jako zgodny ze standardem MP3. Podejście takie ma na celu promowanie różnorodności implementacji programów kodujących i dekodujących dźwięk w standardzie MP3 realizowanych przez różnych producentów. Specyfikacja ISO pełni jedynie rolę bazowego zestawu reguł, określających sposób funkcjonowania standardu tak, aby za pomocą dowolnego koodera można było wygenerować plik odtwarzany przez dowolny dekodery.

2.8 ZALETY I WADY STANDARDU MP3

Niewątpliwie standard kodowania dźwięku MP3 ma wiele zalet. Do najważniejszych należą:

- duży stopień kompresji – stosując kompresję MP3 uzyskujemy plik wynikowy o rozmiarze ok.10 razy mniejszym od oryginału,
- możliwość sterowania stopniem kompresji i tym samym dostosowania jakości dźwięku do indywidualnych potrzeb,
- metoda ta umożliwia uzyskanie sygnałów o stosunkowo dobrej jakości,
- dekompresja wymaga znacznie mniejszej mocy obliczeniowej niż kompresja,



- twórcy standardu bezpłatnie udostępnili kod źródłowy programów kodujących i dekodujących, dzięki czemu standard ten stał się niezwykle popularny.

Warto jednak pamiętać, że MP3 to metoda kompresji stratnej, a tym samym uniemożliwia zrekonstruowanie sygnału oryginalnego. Ocena jakości dźwięku odtworzonego z pliku MP3 jest bardzo indywidualnym doznaniem. Ponieważ algorytm opiera się na matematycznym modelu percepcji słuchowej przeciętnego człowieka, to siłą rzeczy zawsze będzie grupa ludzi, która usłyszy brakujące, wycięte dźwięki. Oczywiście bardzo duże znaczenie będą miały tu parametry dobrane przez twórcę pliku. Osoba nadzorująca proces kompresji MP3 nie ma co prawda bezpośredniego wpływu na współczynnik kompresji lub też na poziom stratności, może jednak ustalać liczbę bitów przypadających na sekundę docelowego zapisu tzw. przepływność. A to przekłada się to bezpośrednio na jakość.

Zadanie 4. Formaty zapisu plików muzycznych, ocena jakości nagrania

Do wykonania zadania jest potrzebny:

- komputer z zainstalowaną kartą dźwiękową i zestawem głośników komputerowych,
- program do przetwarzania dźwięku Audacity, Lame.

W tym ćwiczeniu poznamy jak w prosty sposób przekonwertować plik wav do mp3 i w jaki sposób ocenić jakość nagrania.

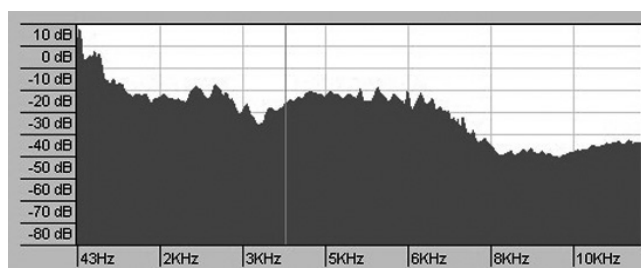
W pierwszej części prowadzący zademonstruje różne nagrania: skompresowane pliki mp3. Pliki są skompresowane z różną przepływnością bitową, a co za tym idzie jakością. Zadaniem słuchaczy jest ocena jakości dźwięku na podstawie odsłuchu. Co wpływa na jakość zarejestrowanego dźwięku?

Na podstawie odsłuchu łatwo jest ocenić jakość dźwięku. Czy jednak można bez słuchania muzyki coś powiedzieć o dźwięku. Patrząc na przebieg sygnału dźwiękowego w czasie – niewiele. Okazuje się jednak, że analiza częstotliwościowa sygnału dźwiękowego daje dużo większe możliwości. Korzystając z narzędzia Analiza [®] Rysuj Widmo z programu Audacity możemy uzyskać reprezentację częstotliwościową sygnału dźwiękowego, czyli tzw. widmo (rys. 29). Na osi poziomej oznaczone są częstotliwości występujące w sygnale dźwiękowym, a na pionowej udział danej częstotliwości w całym sygnale dźwiękowym. Analizując widmo od razu widać w jakim paśmie częstotliwości zawiera się dźwięk i jakie częstotliwości dominują w dźwięku.

Program Audacity umożliwia również tworzenie własnych plików mp3. Potrzebny jest jednak do tego zewnętrzny program Lame. Program Lame uważany jest za jeden z najlepszych, darmowych programów do konwertowania plików dźwiękowych do mp3. Przed wykorzystaniem Lame w Audacity należy go skonfigurować w zakładce Edytuj [®] Ustawienia [®] FormatyPlików [®] Ustawienia_Exportu_dla_MP3. Zakładka ta umożliwia również określenie jakości tworzonych plików mp3.

Program Lame umożliwia także konwersję plików do formatu MP3 z wykorzystaniem linii poleceń. Należy zapoznać się z opcjami wywołania programu z linii poleceń i wykonać konwersję do MP3 tą drogą. Wykorzystanie linii poleceń daje dużo większe możliwości konfiguracji programu.

Słuchacze powinni przekonwertować nagrany wcześniej audycję z formatu wav do formatu mp3 z różną jakością. Należy obejrzeć widma fragmentów tak stworzonych plików. Co można powiedzieć o jakości nagrań oglądając widma sygnałów?



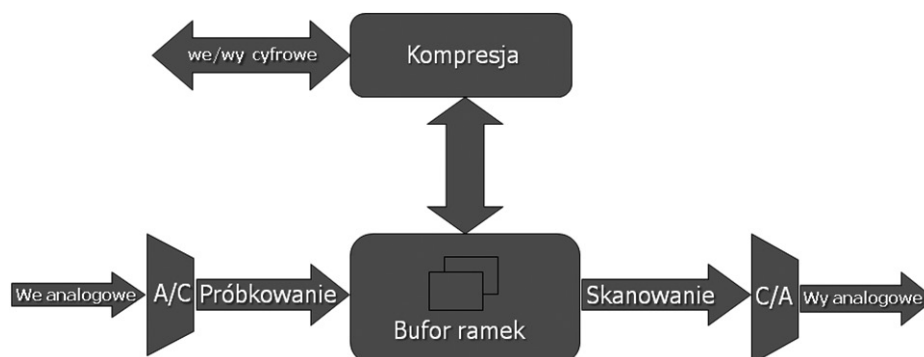
Rysunek 29.
Widmo sygnału dźwiękowego



3. OBRAZY RUCHOME

Cyfrowa obróbka obrazu jest nieodłączną cechą współczesnych systemów multimedialnych. Zanim obraz zostanie poddany cyfrowej obróbce, musi być przekształcony na postać elektryczną (sygnał wizyjny) w przetworniku analizującym (analyzer obrazów), a następnie poddany dyskretyzacji i kwantyzacji. Generalnie operacje przetwarzania sygnału wizyjnego w przykładowym układzie akwizycji obrazu można pogrupować w następujący sposób (rys.30):

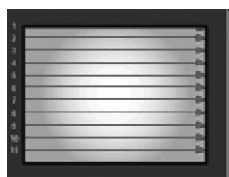
- **Przetwarzanie analogowo-cyfrowe i cyfrowo-analogowe:** Przetwarzanie A/C jest początkowym etapem pozyskiwania obrazu. Przejście z dziedziny analogowej do cyfrowej (i odwrotnie) dokonywane jest przy pomocy wizyjnych przetworników A/C (C/A). Analogowy sygnał wizyjny jest próbkowany w dziedzinie czasu z określoną częstotliwością.
- **Próbkowanie:** Jest to właściwy proces dyskretyzacji obrazu, dokonywany poprzez dwuwymiarowe pobieranie zakodowanych wartości pikseli, niezależny od wcześniejszego próbkowania analogowego sygnału wizyjnego w przetworniku A/C. W zależności od późniejszego wykorzystania strumienia cyfrowej informacji o obrazie (produkcja, transmisja, zapis) stosuje się różne schematy próbkowania cyfrowych strumieni. Pobrane próbki są przechowywane w buforze zawierającym zwykle informacje obrazową z kilku ramek (obrazów). Umieszczenie wartości pikseli w buforze kończy proces pozyskiwania obrazu. Tak zgromadzona informacja stanowi materiał źródłowy dla operacji cyfrowego przetwarzania obrazu.
- **Kompresja obrazu:** Polega na zmniejszeniu ilości informacji o obrazie w celu zapisu na nośniku o ograniczonej pojemności lub transmisji w kanale o określonej szerokości pasma. Kompresja może być bezstratna (zdekodowany materiał jest identyczny z oryginałem), bądź stratna.
- **Skanowanie:** Odczyt zawartości bufora w trybie kolejnoliniowym (bez przeplotu) (rys. 31) lub międzyliniowym (z przeplotem) (rys. 32). Stosowanie przeplotu jest charakterystyczne dla techniki telewizyjnej. Coraz częściej we współczesnych zastosowaniach wideo stosuje się jednak skanowanie progresywne, czyli tryb bez przeplotu.



Rysunek 30.

Proces cyfrowego przetwarzania obrazów

Opisany proces cyfrowego przetwarzania obrazu jest realizowany np. w kamerze cyfrowej. Podczas kręcenia filmu kamerą, najpierw przetwornik (np. matryca CCD) próbkuję światło docierające przez obiektyw i przetwarza je na sygnały elektryczne. Sygnały te są następnie wzmacniane i przesyłane do przetwornika analogowo-cyfrowego (A/C), który nadaje im postać cyfr. Na końcu zaimplementowany w kamerze komputer przetwarza zgromadzone w ten sposób dane cyfrowe, które potem są zapisywane w pamięci jako nowy obraz.



Rysunek 31.

Odczyt zawartości bufora w trybie kolejnoliniowym (skanowanie progresywne)



Rysunek 32.

Odczyt zawartości bufora w trybie międzyliniowym (skanowanie międzyliniowe)

3.1 KOMPUTEROWA EDYCJA OBRAZU

Cyfrowe przetwarzanie obrazów wideo przy pomocy komputera osobistego PC stanowi przykład zaawansowanej formy kształtowania informacji multimedialnych. Do akwizycji materiału wideo stosuje się specjalne karty. Początkowo były to urządzenia typu *frame grabber* umożliwiające przechwyt pojedynczych ramek sekwencji wizyjnej. Wraz z rozwojem technologii układów scalonych, realizujących kompresję, pojawiły się karty akwizycyjne *video capture* umożliwiające przechwycenie ze źródła analogowego (magnetowidu, kamery lub telewizora) sygnału standardów PAL lub NTSC w pełnej rozdzielczości. Skompresowany strumień był zapisywany na szybkim (zwykle SCSI) dysku komputera. Przy pomocy odpowiednich programów można było dokonać edycji zgromadzonego materiału i przeprowadzić stosowny montaż poszczególnych scen. Po zrealizowaniu gotowego wideofilmu można było go poprzez kartę wyprowadzić do zewnętrznego magnetowidu. Taki sposób obróbki sekwencji obrazu nazywany jest **montażem nieliniowym** w odróżnieniu od klasycznego montażu za pomocą dwóch magnetowidów, gdzie sceny materiału końcowego nagrywane są jedna po drugiej. Programy do montażu nieliniowego oferowały coraz większą gamę narzędzi do przetwarzania obrazów oraz towarzyszącego mu dźwięku. Szczególnie wzrastały możliwości w zakresie efektów specjalnych i dodawania dodatkowego tekstu i grafiki. Wraz ze wzrostem mocy obliczeniowej komputerów PC wykonanie nawet najbardziej złożonych trójwymiarowych operacji na obrazie o pełnej rozdzielczości stawało się relatywnie mało czasochłonne a większość efektów mogła być realizowana w czasie rzeczywistym.

Dopiero jednak upowszechnienie się w kamerach amatorskich cyfrowego standardu DV spowodowało znaczące zmiany w konstrukcji kart i programów edycyjnych. Podstawową metodą kompresji stał się standard DV gwarantujący przy stopniu kompresji 5:1 dobrą jakość obrazu o rozdzielczości poziomej 500 linii. Karty zostały wyposażone w interfejsy IEEE 1394 umożliwiające dwukierunkową transmisję skompresowanego sygnału DV. W niektórych kartach pracujących w tym standardzie nawet nie stosuje się gniazd analogowych – podgląd montowanego materiału odbywa się na ekranie monitora komputera, a importowany i eksportowany materiał występuje tylko w postaci cyfrowej. Przy szybkich komputerach kodek DV może być realizowany w wersji programowej. Kariera DVD, jako kolejnego po kasecie DV nośnika cyfrowego, spowodowała pojawienie się kart stosujących kompresję MPEG-2. Jakość sprzętowych koderów MPEG-2 na takich kartach jest już obecnie wystarczająca dla zastosowań amatorskich lub półprofesjonalnych. Przy zastosowaniu odpowiedniego oprogramowania użytkownik może zmontowany materiał wideo nagrać na komputerowej nagrywarce DVD.

Używana w kamerach cyfrowych kompresja DV powoduje jednak pewne, nieznaczne przekłamanie obrazu. Przy nagraniach „naturalnych” pozostają one w zasadzie niezauważalne, ale „sztuczne” obrazy, takie jak: grafika komputerowa, ostre, kontrastowe kształty, jaskrawe kolory, uwidaczniają wyraźnie wady i niekształcenia. Warto jednak zauważyć, że pasmo sygnału nieskompresowanego: $720 \times 576 \times 25$ daje ok. 20MB/s nagrania (kolor jest zapisany 16 bitowo), co jest w większości przypadków nie do przyjęcia. Kompresja DV ma ustalony na stałe współczynnik 5:1, więc redukuje pasmo do 3,7MB/sek. na obraz, do czego dochodzi $2 \times 16 \times 48000 = 1,5 \text{MB/sek.}$ na dźwięk, dając łącznie ok. 5MB/sek., czyli znacznie redukuje ilość danych.

W przypadku kart analogowych proces przechwytywania materiału wideo często nazywany jest **digitalizacją**. W trakcie przeprowadzania tego procesu analogowy materiał wideo zostaje poddany złożonemu procesowi, podczas którego zamieniany jest z postaci analogowej na cyfrową. Istotnym czynnikiem decydującym o jakości digitalizowanego materiału wideo jest rodzaj algorytmu oraz stopień kompresji, im wyższy, tym jakość obrazu jest gorsza. Najczęściej stosowanymi algorytmami kompresji w kartach analogowych są Motion JPEG, MPEG-2. Większość kart analogowych podczas przechwytywania umożliwia regulację podstawowych parametrów obrazu i dźwięku: jasności, kontrastu, nasycenia kolorów, ostrości oraz poziomu głośności dźwięku.

W przypadku kart cyfrowych podczas przechwytywania następuje jedynie cyfrowe kopiowanie skompresowanego materiału wideo z kasety na dysk. Warto dodać, że kamera cyfrowa już w trakcie filmowania,



aby zmniejszyć strumień danych, kompresuje obraz w stopniu 5:1 (kompresja DV). Niezaprzeczną zaletą formatu DV jest zachowanie niezmiennie wysokiej jakości, począwszy od rejestracji, poprzez przechwytywanie, edycję, aż po końcowy zapis. Dla porównania – każde analogowe przechwytywanie, edycja, zapis i kopiowanie stale obniżają jakość.

Zadanie 5. Pozyskiwanie materiału wideo

Sprzęt potrzebny do wykonania zadania:

- kamery wideo pracujące w standardzie miniDV, kamery internetowe, telefon komórkowy z opcją nagrywania,
- cyfrowy aparat fotograficzny.
- oprogramowanie do montażu nieliniowego Pinnacle Studio 9.

Celem tego ćwiczenia jest zapoznanie słuchaczy z technikami przygotowywania materiału filmowego. Jest to jednocześnie pierwszy etap przygotowania własnego filmu.

Cyfrowy format miniDV

Jest to zmniejszona wersja cyfrowego formatu DV. Podstawowa różnica to wielkość kasety, dzięki której zmniejszono rozmiar kamer. Jednak wskutek niewielkich rozmiarów nośnika skraca się długość taśmy. W trybie Standard Play można nagrać 60 minut nagrania, natomiast w trybie Long Play aż 120 minut. Ogólnie jakość obrazu i dźwięku jest bardzo wysoka. Dźwięk można zapisywać na 4 lub 2 kanałach. Jest porównywalny z jakością CD. Format miniDV zapewnia ponad 500 linii rozdzielczości poziomej. Kopiowanie odbywa się bezstratnie tzn. można przegrywać taśmę wielokrotnie bez pogorszenia się jakości obrazu. Do przesyłania obrazów do komputera służy wyjście DV (kompatybilne ze złączem IEEE1394, tzw. FireWire).

Na wstępie należy się zapoznać ze sprzętem. Szczególną uwagę należy zwrócić na sposób obsługi i funkcje kamer cyfrowych. Na uwagę zasługuje system **3xCCD**, czyli potrójna matryca CCD, umożliwiająca rejestrowanie najwyższej jakości obrazów w doskonałej rozdzielczości. Podstawą technologii zawartej w przetwornikach obrazu 3CCD jest pryzmat, który rozszczepia światło na trzy podstawowe kolory RGB (niebieski, czerwony i zielony). W ten sposób każdy z przetworników analizuje osobno odpowiednią barwę, wpływając na dokładność nagrywanego obrazu. Jest to więc analiza trójprzetwornikowa.

Na etapie przygotowania materiału filmowego gromadzimy różne media: pliki, które będą później elementami montowanego filmu: grafiki, dźwięki i klipy filmowe. Na tym etapie najpierw należy przygotować **własną koncepcją krótkiego filmu** do montażu. Film będzie montowany z różnych elementów dlatego należy przygotować: materiał filmowy nagrany różnymi kamerami, własne komentarze nagrane przy użyciu oddzielnych mikrofonów, zdjęcia wykonane aparatem cyfrowym. Wszystkie te elementy mogą być zapisywane w różnych formatach. Później będą one łączone w jeden film.

Każdy słuchacz (lub zespół) powinien na tym etapie mieć:

- przygotowaną koncepcję własnego krótkiego filmu do montażu (bez użycia komputera należy zastanowić się, zapisać co się chce przedstawić),
- skompletowane klipy filmowe nagrane kamerami pracującymi w standardzie miniDV, kamerami internetowymi i ewentualnie nagrane przy użyciu telefonu komórkowego z opcją rejestracji wideo,
- zdjęcia, które mają być później ilustracją do filmu (np. plansze tytułowe, przerywniki),
- muzyka mogąca służyć za podkład do scen w filmie.

Zgrywanie sekwencji wizyjnej

Pierwszą czynnością, od której rozpoczyna się każdy montaż nieliniowy, jest przeniesienie zapisanego materiału wideo z kasety na twardy dysk komputera określane potocznie przechwytywaniem. Celem tego ćwiczenia jest zapoznanie słuchaczy z metodami zgrywania sekwencji wizyjnej na dysk komputera.

Pakiet Pinnacle Studio 9 umożliwia zgrywanie materiału filmowego zarówno bezpośrednio z cyfrowej kamery DV jak i ze źródła analogowego za pomocą analogowej karty przechwytyującej. Ponieważ dysponujemy nagraniami cyfrowymi będziemy korzystać z pierwszej opcji. Zazwyczaj istnieje możliwość zapisania materiału w postaci pliku AVI (format DV) jak i zgranie materiału bezpośrednio do formatu MPEG lub WMV. W naszym przypadku zgrywamy materiał do pliku .avi z kompresją DV. Zwykle



do zgrywania materiału lepiej jest wykorzystywać magistralę FireWire niż USB. FireWire jest powszechnie używany do łączenia kamer wideo i urządzeń pamięci masowej. Stosuje się go zamiast popularniejszego USB z powodu większej szybkości transmisji (prędkość nie zależy od wielkości plików jak przy USB). Nie ma również konieczności wysyłania sygnałów potwierdzających aktywność urządzenia po drugiej stronie (co czyni USB nieefektywnym dla profesjonalnej obróbki wideo).

Dużym problemem przy przechwytywaniu jest bariera dopuszczalnej długości pliku istniejąca w systemie Windows. W przypadku Windows 95 wynosi ona 2 GB, w Windows 98/Me wynosi 4 GB, a w przypadku Windows NT4.0/2000/XP dla dysków sformatowanych w formacie NTFS wynosi 12 TB (czyli można uznać, że w praktyce nie istnieje). Problemem jest natomiast to, że pakiety programowe, jak Pinnacle Studio 9, nie zawsze dobrze radzą sobie z dużymi plikami. Dużo zależy też od konfiguracji komputera. Dlatego wygodnie jest na etapie przechwytywania podzielić plik wideo na mniejsze części. Takie pliki jest dużo łatwiej przetwarzać i przechowywać.

Wygodną metodą przeszukiwania jest tzw. skanowanie taśmy, podczas którego automatycznie zostają wyszukane wszystkie zmiany kodu czasowego powstałe na taśmie w wyniku zatrzymania procesu nagrywania (przez naciśnięcie przycisku Stop lub Pause). W rezultacie uzyskujemy gotową listę cięć, według której można automatycznie przechwycić materiał wideo podzielony na poszczególne ujęcia. Taki sposób przechwytywania umożliwiają również pakiet Pinnacle Studio 9.

Na etapie zgrywania sekwencji wizyjnej słuchacze powinni:

- zapoznać się z metodami łączenia poszczególnych urządzeń (kamery, aparat fotograficzny, cyfrowy aparat fotograficzny, mikrofony i wzmacniacze mikrofonowe z komputerem,
- zapoznać się z metodami, możliwościami, magistralami umożliwiającymi skopiowanie materiałów multimedialnych na dysk komputera (FireWire i USB),
- zapoznać się z zasadą działania i podstawowymi funkcjami programu i pakietu oprogramowania Pinnacle Studio 9,
- przygotować materiały do montażu
- materiały własne przechwycone z kamery przy użyciu programu Pinnacle Studio 9,
 - materiały własne przechwycone z kamery internetowej lub telefonu komórkowego,
 - zdjęcia cyfrowe,
 - gotowe przykładowe sekwencje wizyjne,
- przygotować dodatkowe materiały graficzne (np. plansze, napisy),
- jeszcze raz przemyśleć koncepcję własnego krótkiego filmu do montażu (bez użycia komputera, należy zastanowić się, co chcemy zamieścić w filmie, które materiały będą użyte, a które nie).

3.2 STANDARDY KODOWANIA SEKWENCJI WIDEO MPEG1 I MPEG2

Skrót MPEG pochodzi od nazwy grupy roboczej *Moving Picture Experts Group*, opracowującej standardy kompresji audio-video na potrzeby międzynarodowej organizacji normalizacyjnej ISO.

Potrzeba kompresji audio pojawiła się podczas prób zapisania analogowego sygnału wizji i fonii w postaci cyfrowej. Przykładowo popularny system przesyłania sygnału telewizyjnego PAL przekazuje obraz telewizyjny składający się z 25 klatek na sekundę o rozdzielczości 704 na 576 i każdy punkt jest opisany 24-bitową głębią kolorów. Próba zapisanie takiego obrazu cyfrowo powoduje, że każda sekunda obrazu zajmuje około 30 megabajtów. Na początku lat 90. XX wieku opracowany został standard kompresji **MPEG-1**, który umożliwia przesyłanie obrazu audio-video z przepustowością 1,5 Mb/s przy rozdzielczości ekranu 352x240 lub 352x288. Standard ten umożliwił stworzenie cyfrowego zapisu audio-video Video CD, którego jakość była porównywalna do standardu VHS. Niskie rozdzielczości, jakie oferuje standard MPEG-1, powodują zniekształcenia obrazu przy pełnoekranowym wyświetlaniu, przez co nie nadaje się on do przesyłania obrazów dobrej jakości, a takie wymagania stawia między innymi telewizja kablowa i satelitarna. To właśnie na potrzeby telewizji cyfrowej w połowie lat 90. XX wieku powstał standard **MPEG-2**, który umożliwia przesyłanie obrazów o znacznie większych rozdzielczościach, aż do 1920 x 1152 punktów, i przepustowości między 3 a 100 Mb/s. Opublikowanie standardu MPEG-2 stało się przełomowym wydarzeniem dla techniki telewizyjnej. Standard ten określający metodę kompresji i kodowania sygnału wizyjnego, fonii i danych dodatkowych otworzył drogę do opracowania i wdrożenia cyfrowych standardów emisji programów telewizyjnych. Jest to pierwszy cyfrowy standard opracowany pod kątem zastosowania w telewizji programowej. Wcześniejsze standardy: JPEG,



MPEG-1 były przeznaczone do innych celów i znalazły tylko ograniczone zastosowanie w telewizji. Standard **MPEG-4**, opracowany w 1999 roku, przystosowany został głównie do kompresji danych strumieniowych (np. wideokonferencje), dlatego ma zaimplementowane funkcje ochronne przed błędami przesyłu. MPEG-4 oferuje najwyższy stopień kompresji z całej rodziny standardów MPEG.

We wszystkich standardach MPEG stosowana jest stratna metoda kompresji wizji, tzn. sygnał po dekompresji różni się od pierwotnego. Wykorzystuje się przy tym:

- właściwości oka ludzkiego – oko ludzkie nie jest idealnym przetwornikiem, nie jest w stanie dostrzec wszystkich detali w transmitowanych obrazach, a więc z obrazów tych można usunąć część informacji nie powodując przy tym pogorszenia subiektywnej oceny jakości obrazu;
- korelację przestrzenną (wewnątrzobrazową) – w większości obrazów występują jednolite lub nieznacznie tylko różniące się fragmenty, dzięki zastosowaniu dyskretnej transformaty kosinusowej DCT (ang. *Discrete Cosine Transform*) obszary takie mogą być efektywnie zakodowane;
- korelację czasową – kolejne obrazy w sekwencji składającej się na obraz telewizyjny z reguły niewiele od siebie się różnią, na podstawie bieżącego obrazu można z dobrym przybliżeniem wyznaczyć następny obraz w sekwencji jeśli zostaną wykryte przesuwające się obiekty i prawidłowo zostanie opisany ich ruch, metoda ta nazywana jest także metodą kompensacji ruchu;
- właściwości statystyczne sygnału – w transmitowanym sygnale pewne symbole pojawiają się częściej, a inne rzadziej, zastosowanie do ich kodowania słów o różnej długości – krótszych słów dla symboli częściej się pojawiających, dłuższych dla pozostałych umożliwia zmniejszenie przepływności binarnej sygnału.

Aby przedstawić sposób kompresji obrazów, konieczne jest zrozumienie podstawowych faktów związanych z obrazem kolorowym. Próby opisanie obrazu kolorowego doprowadziły do powstania modelu RGB, który opisuje nasycenie trzech barw (czerwonej, zielonej i niebieskiej) dla każdego punktu obrazu, w urządzeniach, które emitują światło. Odpowiednie dobranie poszczególnych parametrów umożliwia generowanie wielu barw widzianych przez ludzkie oko. Model ten ma pewne niedoskonałości (np. problem z uzyskaniem czystej bieli), ale mimo wad jest on wystarczający do reprezentowania obrazu kolorowego.

Na potrzeby przesyłania kolorowego sygnału telewizyjnego zostały opracowane modele, w których obraz jest kodowany jako jedna składowa **luminancji** (jasności) oraz dwie składowe **chrominancji** (koloru) – $Y C_R C_B$. Poszczególne składowe są obliczane na podstawie wartości dla poszczególnych kolorów RGB mnożonych przez różne współczynniki (patrz wzory niżej).

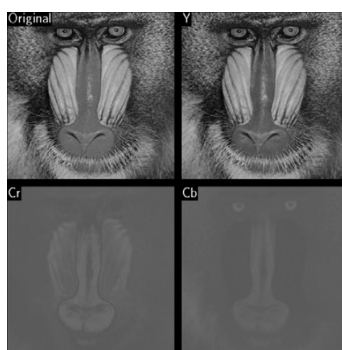
$$Y = 0.299(R - G) + G + 0.114(B - G)$$

$$C_b = 0.564(B - Y);$$

$$C_r = 0.713(R - Y)$$

gdzie Y to luminancja, a C_b , C_r to sygnały chrominancji.

Specyfika ludzkiego wzroku, dla którego bardziej znaczące jest informacja o jasności punktu niż o kolorze, umożliwia pominięcie niektórych informacji zawartych w modelu $Y C_R C_B$, dzięki temu ograniczając ilość informacji o chrominancji można uzyskać obraz o zbliżonej jakości, ale z mniejszą liczbą informacji potrzebnych do wygenerowania tego obrazu. Obraz rozłożony na składowe chrominancji i luminancji jest przedstawiony na rys. 33.

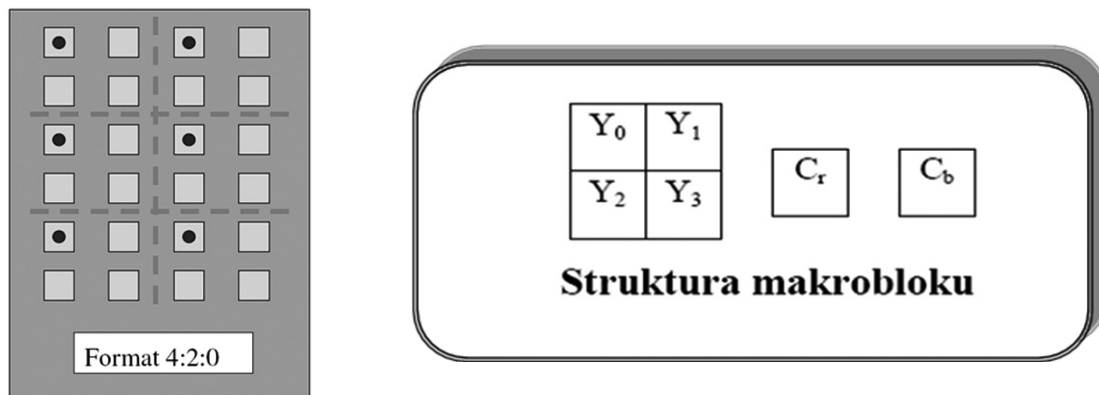


Rysunek 33.

Składowe luminancji Y i chrominancji $C_R C_B$ obrazu kolorowego

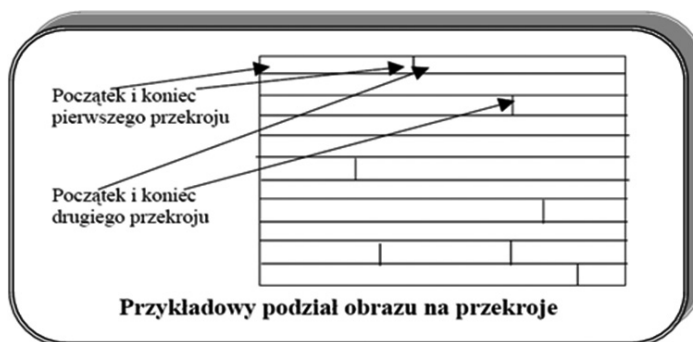


Dla standardu MPEG-1 przyjęto strukturę próbkowania obrazu 4:2:0 (na 4 próbki luminancji przypadają dwie próbki chrominancji w jednej linii, oraz 0 próbek chrominancji w kolejnej linii). Taka struktura próbkowania przedstawiona jest na rys. 34a. Wartości brakujących próbek chrominancji wyznacza się poprzez interpolację (obliczenie wartości przybliżonych). Elementarną porcją obrazu kodowaną przez koder MPEG nazywa się **makroblokiem**. Makroblok to tablica 16x16 pikseli złożona z próbek luminancji i odpowiadające im tablice chrominancji C_r i C_b o rozmiarze 8x8 pikseli (rys. 34b).



Rysunek 34. Format próbkowania obrazu 4:2:0 stosowane w standardzie MPEG-1 oraz struktura makrobloku

W następnym kroku makrobloki są składane w tzw. **przekroje**, które mogą zawierać różną liczbę makrobloków (rys. 35). Niezależne kompresowanie każdego przekroju zabezpiecza przed rozszerzaniem się ewentualnych błędów na cały obraz.



Rysunek 35. Podział obrazu na przekroje

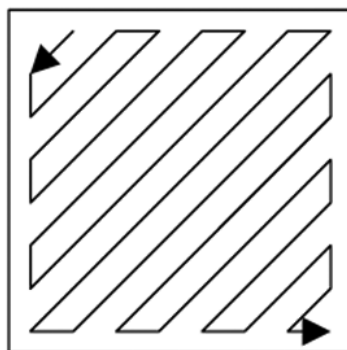
W dalszych etapach kodowania MPEG wykorzystuje się transformację kosinusową, kwantyzację, kodowanie Huffmana. Procesy te zostały wyjaśnione przy okazji omawiania kompresji MP3.

3.2.1 PRZEBIEG PROCESU KODOWANIA W STANDARDACH MPEG-1 I MPEG-2

Kompresja MPEG wykorzystuje to, że obraz wideo ulega stopniowym zmianom, a część tego obrazu, w jakimś okresie czasu, często jest niezmienna. Poprzez zachowanie informacji raz na około 0,5 sek., o pełnych pojedynczych klatkach (**ramkach**) filmu oraz informacji o zmianach jakie następują w ciągu tego czasu możliwe jest znaczne ograniczenie wielkości przekazu wideo. Dzięki zastosowaniu specjalnych algorytmów możliwe jest generowanie ramek na podstawie informacji o zmianach oraz stanie poprzednich i następnych ramek.

Wyróżniamy trzy typy ramek **I**, **P** i **B**. Typ pierwszy **I**, to ramki zawierające pełną informację o obrazie. Ramki typu **P** są tworzone na podstawie obrazu poprzedniej ramki typu **I** lub **P**. Zawierają one informacje o elementach zmiennych w sekwencji wideo oraz o kierunku przesunięcia zmian (ruchu). Ramki typu **B** są tworzone na podstawie dwóch obrazów poprzedniego typu **I** lub **P** oraz następnego typu **I** lub **P**, według zasad stosowanych przy ramach typu **P**.

Obrazy **typu I** są kodowane w sposób następujący: cały obraz jest dzielony na bloki o wymiarach 8x8 pikseli (sygnał luminancji niezależnie od sygnału chrominancji), a następnie dla każdego bloku (niezależnie od pozostałych bloków) jest wyznaczana jego transformata kosinusowa. Transformata kosinusowa jest przekształceniem odwracalnym i z jej współczynników można odtworzyć pierwotny blok. Współczynniki transformaty kosinusowej zawierają informację o tym, jak szybko zmieniają się wartości próbek w przetwarzanym bloku. W większości obrazów występują obszary stałe, wartości próbek bloków należących do takich obszarów niewiele różnią się od siebie. Natomiast współczynniki transformaty kosinusowej mają znacznie różniące się wartości – tylko współczynnik odpowiadający składowej stałej i sąsiednie współczynniki odpowiadające niskim częstotliwościom mają duże wartości, pozostałe odpowiadające wyższym częstotliwościom są równe lub bliskie zeru. Z badań nad właściwościami wzroku ludzkiego wiadomo, że oko jest znacznie bardziej wrażliwe na składowe niskoczęstotliwościowe obrazu niż na składowe wysokoczęstotliwościowe. Jeśli więc będziemy kwantować współczynniki transformaty kosinusowej, to współczynniki niskoczęstotliwościowe powinny być kwantowane dokładniej, z dużą liczbą bitów, a współczynniki wysokoczęstotliwościowe mniej dokładnie. W wyniku takiej kwantyzacji większość współczynników wysokoczęstotliwościowych, które były bliskie zeru, przyjmie zerowe wartości. Na etapie kwantyzacji następuje nieodwracalna utrata części informacji o obrazie, ze skwantowanych współczynników zostaną odtworzone próbki różniące się od pierwotnych. Jeśli jednak poszczególne współczynniki będą wystarczająco dokładnie skwantowane (zostanie im przypisana dostatecznie duża liczba bitów), to różnice pomiędzy pierwotnym, a odtworzonym blokiem obrazu będą niedostrzegalne przez człowieka. Współczynniki transformaty kosinusowej są wybierane według określonego porządku (wybieranie zygzakowe) (rys. 36). Wybieranie zygzakowe jednocześnie szereguje współczynniki od najbardziej istotnych do najmniej ważnych z punktu widzenia kompresji obrazu.



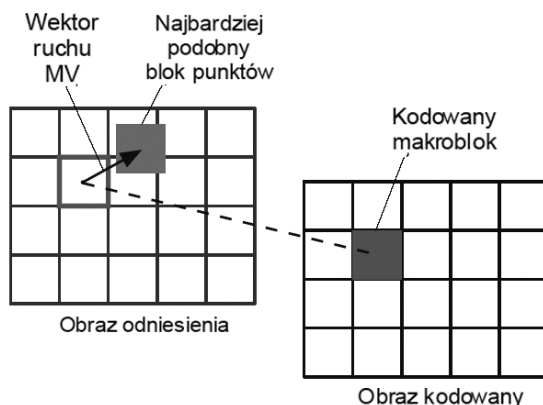
Rysunek 36.

Sposób wybierania współczynników transformaty DCT – „zygzak”

Współczynnik DC (średnia wartości pikseli w bloku, której odpowiada pierwszy współczynnik transformaty kosinusowej) jest kodowany różnicowo względem predykcji czyli współczynnika DC z poprzedniego bloku. Z pozostałej części macierzy współczynników DCT tworzona jest sekwencja współczynników według porządku „zygzak”. Następnie po kwantyzacji współczynniki są kodowane kodem Huffmana.

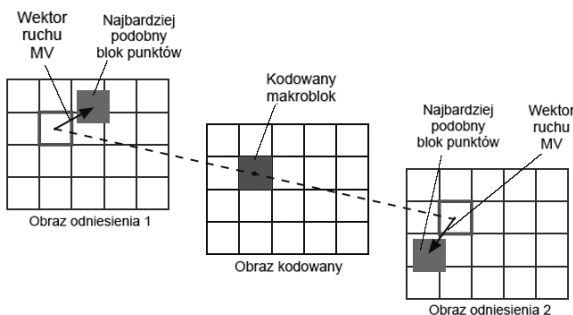
Przy kodowaniu obrazów **typu P** korzysta się z podobieństwa obrazów w sekwencji. Dlatego koduje się nie poszczególne obrazy, ale różnice pomiędzy kolejnymi obrazami. Dwa kolejne obrazy są odejmowane od siebie i ich różnicę koduje się tak, jak omówione wcześniej obrazy typu I, wykorzystując transformatę kosinusową i kwantyzację jej współczynników. Ponieważ jest to kompresja stratna więc transmitowana informacja różni się od wyniku odejmowania kolejnych obrazów. Ta różnica mogłaby spowodować rozbieżności pomiędzy procesami kodowania i dekodowania obrazów. Żeby tego uniknąć w koderze odtwarza się poprzez odwrotną transformatę kosinusową i odwrotną kwantyzację dokładnie taki sam obraz jaki powstanie w dekodecie – obraz ten nazywany jest **obrazem odniesienia** (rys. 37). Kodowaniu poddawana jest więc różnica pomiędzy bieżącym obrazem a obrazem odniesienia. Przedstawiony schemat kodowania jest skuteczny, jeśli kolejne obrazy w sekwencji niewiele różnią się od siebie, a więc gdy transmitowana scena jest prawie statyczna. Jeśli jednak występują w niej ruchome elementy, to metoda kompresji oparta na kodowaniu różnic pomiędzy tymi obrazami przestaje być efektywna.





Rysunek 37.
Sposób kodowania obrazów typu P

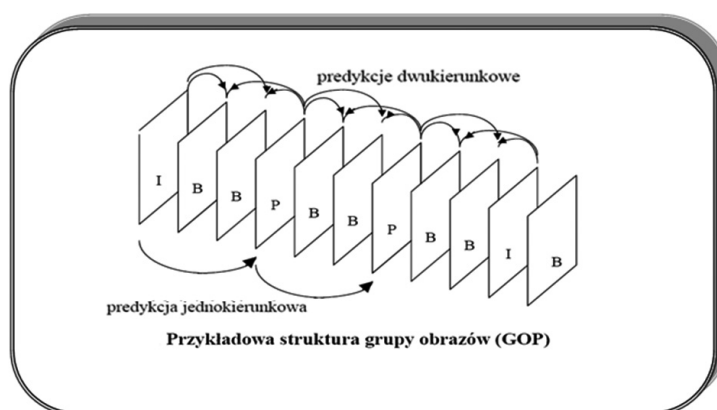
Obrazy typu B kodowane są podobnie jak obrazy typu P, z tym że jednocześnie wykorzystuje się dwa obrazy odniesienia – wcześniejszy i późniejszy. Prognoza kodowanego obrazu tworzona jest przez uśrednienie informacji pochodzących od obu obrazów odniesienia, dlatego też ten typ obrazu nazywany jest **obrazem interpolowanym** (rys. 38). Metoda kodowania obrazów typu B wymaga, aby dekodery wcześniej znał oba obrazy odniesienia. Obrazy odniesienia muszą być wysłane wcześniej niż zakodowane obrazy typu B. Bez nich niemożliwe jest rozkodowanie.



Rysunek 38.
Sposób kodowania obrazów typu B

Poszczególne typy obrazów I, P i B stosowane w standardzie MPEG zapewniają różne stopnie kompresji. Najmniejszy stopień kompresji umożliwiają obrazy typu I, gdyż zawierają pełną informację o obrazie. Obrazy te muszą być jednakże dość często transmitowane, gdyż odbiornik może rozpocząć dekodowanie sekwencji tylko od tego typu obrazu. Obrazy typu P zapewniają większy stopień kompresji, przy ich kodowaniu powstają jednak pewne zniekształcenia spowodowane skończoną dokładnością obliczeń, które kumulują się i dlatego liczba kolejno następujących po sobie obrazów typu P nie może być zbyt duża. Jeszcze większy stopień kompresji umożliwiają obrazy typu B. W tym przypadku także występują wspomniane wcześniej zniekształcenia, a ponadto przy transmisji obrazów tego typu musi zostać zmieniona kolejność nadawania obrazów – najpierw muszą być wysłane oba obrazy odniesienia, a dopiero potem obraz typu B. Wymaga to wyposażenia dekodera w odpowiednio dużą pamięć.

Liczba obrazów poszczególnych typów jak i długość całej **grupy obrazów** nie jest określona w standardzie MPEG i może być różna w zależności od wymagań stawianych danemu systemowi transmisji. Struktura grupy obrazów nie musi być stała w czasie całej transmisji i może się zmieniać w zależności od aktualnej treści przekazywanego programu. Przykładowa struktura grupy obrazów jest przedstawiona na rys. 39.



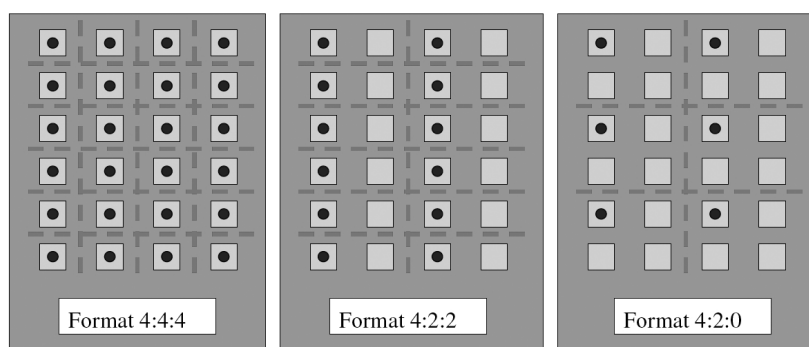
Rysunek 39.

Struktura grupy obrazów

3.2.2 RÓŻNICE POMIĘDZY STANDARDAMI MPEG-1 I MPEG-2

Standard MPEG-2 został opracowany z myślą zastosowania w cyfrowych systemach telewizyjnych, do obsługi których MPEG-1 się nie nadawał. MPEG-2 nie jest związany z żadnym wykorzystywanym obecnie analogowym standardem telewizyjnym. Zgodnie z tym standardem można transmitować zarówno programy telewizyjne zrealizowane w europejskim standardzie 625 linii/50Hz, jak i amerykańskim 525 linii/60Hz. W standardzie MPEG-2 można także przysyłać programy zrealizowane w standardach wysokiej rozdzielczości HDTV. Standard MPEG-2 dopuszcza stosowanie zarówno międzyliniowego, jak i kolejnoliniowego wybierania obrazu, format ekranu może być 4:3 lub 16:9. Dla zapewnienia możliwości odtwarzania obrazów panoramicznych na zwykłych odbiornikach przesyłana jest w nim informacja o tym, która część obrazu powinna być w danej chwili wyświetlana. Standard MPEG-2 dopuszcza także podział sygnału na hierarchicznie zorganizowane warstwy, które umożliwiają np. jednoczesną transmisję jednego programu telewizyjnego w standardowej i dużej rozdzielczości. Dzięki kompresji sygnału wizyjnego standard MPEG-2 umożliwia kilkukrotne zwiększenie liczby nadawanych programów w dostępnych obecnie sieciach transmisji programów telewizyjnych. Wprowadzono również nowe elementy, jak przetwarzanie dźwięku w pięciokanałowym systemie przestrzennym *surround*, wielojęzyczne napisy oraz znaczniki np. do znajdowania określonych fragmentów filmu.

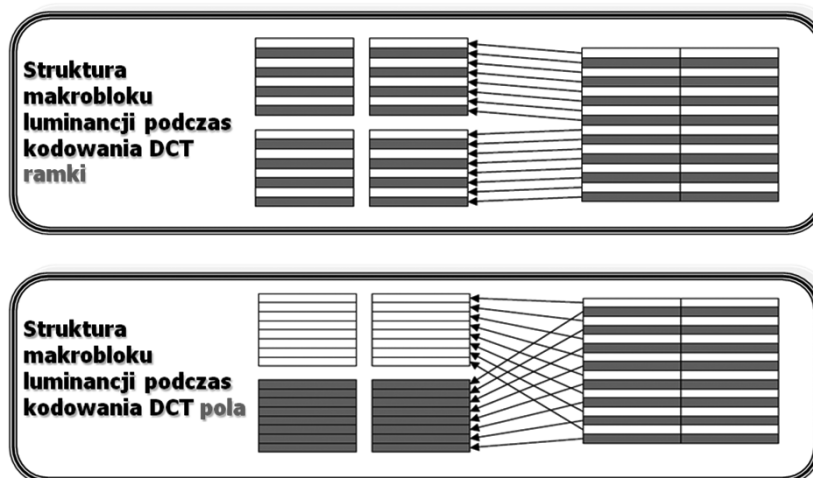
Standard MPEG-2 wykorzystuje strukturę próbkowania 4:2:0 lub 4:2:2 (rys. 40). W strukturze próbkowania 4:2:0 następuje zmniejszenie rozdzielczości sygnałów chrominancji w obu kierunkach: pionowym i poziomym. Makrobok w standardzie 4:2:0 składa się z 4 bloków luminancji, jednego bloku chrominancji C_r i jednego bloku chrominancji C_b . Jeden blok zawiera 8x8 pikseli. Natomiast w strukturze próbkowania 4:2:2 liczba próbek chrominancji w kierunku pionowym nie ulega zmianie, zmniejsza się natomiast liczbę tych próbek w kierunku poziomym. Makrobok w standardzie 4:2:2 składa się z 4 bloków luminancji, 2 bloków chrominancji C_r i dwóch bloków chrominancji C_b . Dla porównania pokazano również strukturę próbkowania 4:4:4, gdzie na każdą próbkę sygnału luminancji Y przypada po jednej próbce sygnału chrominancji C_r i C_b .



Rysunek 40.

Formaty próbkowania obrazu 4:2:0 i 4:2:2 stosowane w standardzie MPEG-2

W przypadku wybierania międzyliniowego, ramka sygnału wizyjnego składa się z dwu pól. Jako obraz kodowana może być ramka (każdy blok powinien być złożony naprzemiennie z linii dwóch pól) lub pole (każdy blok powinien być złożony tylko z linii jednego z dwóch pól). Specyfikacja MPEG-2 umożliwia kodowanie ramki jako jednego obrazu oraz kodowanie niezależnie dwóch obrazów (kodowanie pola) – rys. 41.



Rysunek 41.

Struktura makrobloku luminancji podczas kodowania ramki i pola

Zmiana typu kodowania może następować w każdej kolejnej ramce. Kodowanie ramki daje lepsze rezultaty, gdy w sekwencji obrazów jest mało ruchu – półobrazy dobrze uzupełniają się przestrzennie. Kodowanie pola jest lepsze przy dużym natężeniu ruchu w sekwencji obrazów – korelacja przestrzenna półobrazów jest załkóna i lepiej jest użyć kompensacji ruchu.

Zadanie 6. Edycja sekwencji wizyjnej

Sprzęt potrzebny do wykonania zadania:

- komputer,
- oprogramowanie do montażu nieliniowego Pinnacle Studio 9.

Edycja sekwencji wizyjnej jest bez wątpienia najważniejszym procesem każdego montażu materiału wideo. To właśnie dzięki niej mamy możliwość uporządkowania często chaotycznie zarejestrowanych ujęć. Wykorzystując możliwości montażu nieliniowego, możemy znacznie uatrakcyjnić i ożywić niezbyt ciekawe nagrania. Podstawowymi zabiegami edycyjnymi są precyzyjne docinanie, kopiowanie i układanie poszczególnych ujęć w żądanej kolejności. Ponadto możliwe jest tworzenie stop-klatek, regulacja tempa oraz kierunku odtwarzania klipu wideo. Wszystkie programy edycyjne wyposażone są w ścieżki wideo i audio, oś lub linię czasu oraz podstawowe narzędzia edycyjne. Nasze edytowane ujęcia umieszczane są na poszczególnych ścieżkach wideo i audio. Podstawową operacją przeprowadzaną podczas edycji jest cięcie, które ma wyodrębnić poszczególne ujęcia z materiału wideo. Do tego celu używamy zazwyczaj wirtualnych "nożyczek". Precyzyjne dobieranie początku i końca ujęcia możliwe jest również dzięki zastosowaniu znaczników wejścia i wyjścia. W niektórych programach dostępna jest opcja trymowania, która w precyzyjny sposób kontroluje i jednocześnie obserwuje proces przycinania ujęcia. Stop-klatki w montowanym materiale wideo można uzyskać na dwa sposoby. Pierwszym z nich jest zaimportowanie do programu montażowego klatki w postaci pliku graficznego np. BMP, TIFF, JPEG, jak również ustawienie jej bezpośrednio z ujęcia umieszczonego na ścieżce wideo. Zmiana tempa odtwarzanego materiału wideo wpływa na zwolnienie lub przyśpieszenie akcji filmu, natomiast zmianę kierunku odtwarzania możemy zastosować w celu uzyskania efektu specjalnego. Oprócz możliwości odpowiedniego przycięcia sekwencji na tym etapie możemy rozjaśnić bądź przyciemnić sekwencję wideo, lub też zastosować różnego rodzaju filtry.

Wszystkie programy do obróbki wideo, w tym Pinnacle Studio 9, na etapie edycji oferują dwa tryby pracy: Storyboard i Timeline. Tryb **Storyboard** jest najszybszą i najprostszą metodą dodawania

klipów do tworzonego właśnie filmu, poprzez zastosowanie przenoszenia poszczególnych sekwencji. W trybie tym każdy obrazek reprezentuje pojedyncze zdarzenie w edytowanym filmie. Istnieje możliwość przenoszenia żądanych klipów filmowych i umieszczania ich w założonym miejscu. Dodatkowo pomiędzy klipami można dołączyć różne efekty przejścia. Tryb **Timeline** umożliwia dokładne dopasowywanie efektów i edycję materiału. W trybie tym pokazywane jest dokładnie chronologia pojawiania się poszczególnych elementów filmu na poszczególnych ścieżkach. Klipy filmowe w tym trybie są mogą być przenoszone z bibliotek na odpowiednie ścieżki i są wyświetlane jako krótkie sekwencje.

Elementy edycji klipów filmowych obejmują:

- dodawanie klipu filmowego do ścieżki,
- przycinanie klipu filmowego,
- podział klipu filmowego na segmenty,
- podział na sceny,
- ekstrakcja materiału,
- korzystanie z filtrów.

a) Efekty specjalne

Obecnie każdy program edycyjny ma różnego rodzaju efekty specjalne, które możemy podzielić na: efekty przejść (miksy), filtry obrazu, kluczowanie obrazu, nakładanie obrazu oraz wprowadzanie napisów. Z efektów przejść korzystamy w sytuacji, gdy zamierzamy w atrakcyjny sposób połączyć ze sobą dwa różne ujęcia. Generalnie, efekty przejść można podzielić na dwuwymiarowe (2D) i trójwymiarowe (3D). Najprostszymi efektami dwuwymiarowymi są przenikania obrazu (miksy) oraz tzw. kurtynki, natomiast efekty trójwymiarowe najczęściej kojarzone są z efektem odwijanej kartki lub transformacji obrazu w trójwymiarową bryłę. Rodzaj oraz liczba efektów przejść zależne są od klasy i typu programu edycyjnego.

Filtry obrazu mogą spełniać rozmaite funkcje. Pierwszą z nich jest korekcja parametrów obrazu. Do tego celu przeznaczone są filtry regulujące balans kolorów (RGB), jasność, kontrast, nasycenie itp. Drugą z nich jest wprowadzanie prostych efektów, takich jak odwracanie obrazu w pionie i poziomie, docinanie itp. Natomiast trzecią z nich jest wprowadzanie efektów specjalnych: rozmycia, wyostrenia, płaskorzeźby, kalejdoskopu, fali, deformacji, efektów świetlnych itp.

Kolejny efekt **kluczowania** w uproszczeniu polega na usunięciu z obrazu wideo pewnych obszarów charakteryzujących się wspólnymi cechami (np. kolorem, jasnością itp.) i nadaniu im przezroczystości. Dzięki kluczowaniu uzyskujemy możliwość nakładania na siebie poszczególnych warstw z kolejnymi obrazami wideo lub grafiką. Najczęściej stosowanymi rodzajami kluczy są:

- *color key* – usuwanie obszarów o określonym kolorze,
- *chroma key* – usuwanie obszarów o określonym kolorze wraz z jego odcieniami (działa dużo precyzyjniej od *color key*, ponieważ uwzględnia poziom nasycenia koloru),
- *luma key* – usuwanie obszarów jedynie o określonym poziomie jasności,
- *blue screen* – usuwanie niebieskich obszarów (często zamiast koloru niebieskiego używa się koloru zielonego oraz klucza *green screen*) – są to najczęściej stosowane klucze używane do usunięcia tła za filmowaną postacią, np. spikerem w wiadomościach lub prognozie pogody,
- *alpha channel* – usuwanie obszarów zdefiniowanych specjalną maską zapisaną w dodatkowym kanale alfa. Niestety często się zdarza, że efekty dalekie są od ideału. Wynika to z trudności w uzyskaniu jednolitych obszarów obrazu, które mogłyby być precyzyjnie usunięte. Profesjonalne rozwiązania posługują się specjalnie przygotowywanymi maskami umożliwiającymi bardzo precyzyjne kluczowanie.

Pakiet Pinnacle Studio 9 umożliwia miksowanie i zastosowanie zestawu dostępnych efektów podczas edycji materiału. Wszystko to umożliwia uzyskanie profesjonalnego wyglądu tworzonego filmu.

b) Nakładanie scen

Kolejnym efektem jest nakładanie obrazu. Mamy tu do czynienia z efektami powstałymi w wyniku nałożenia na obraz wideo jednego lub więcej często zdeformowanych obrazów pochodzących z plików wideo lub plików graficznych. Dobrym tego przykładem mogą być efekty typu obraz w obrazie (ang. **PIP**



– *Picture In Picture*). Deformacje mogą dotyczyć zarówno wielkości, jak i kształtu nakładanego obrazu, który dodatkowo może być statyczny lub ruchomy. Wprowadzanie napisów jest efektem polegającym na nałożeniu na obraz wideo przede wszystkim napisów, ale również grafik lub małych animacji. Nakładanie napisów jest możliwe dzięki towarzyszącemu mu nieodłącznie efektowi kluczowania (najczęściej alpha channel). Napisy mogą być zarówno statyczne, jak i dynamiczne. Większość programów do montażu wideo ma wbudowane proste funkcje do tworzenia napisów, jednak oferują one jedynie podstawowe możliwości.

Bardzo ważnym elementem każdego filmu jest ekran tytułowy oraz ekran z napisami końcowymi. Aby stworzyć napisy, które mają być potem wykorzystane przy sekwencji początkowej lub końcowej, należy wybrać ramkę, do której dodawać chcemy napisy tytułowe. Następnie tworzymy napis tytułowy. Do tekstu można dodać efekty, np. napisy przewijane, napisy zanikające, inne efekty animacji.

Na etapie edycji wideo słuchacze powinni dokonać montażu wideo z zastosowaniem kroków przedstawionych w punkcie a) i b) niniejszej instrukcji.

4 ODTWARZANIE TREŚCI MULTIMEDIALNYCH

4.1 TELEWIZJA ANALOGOWA I CYFROWA

Prace nad systemem telewizji kolorowej, rozpoczęły się w połowie lat 50. XX wieku w Stanach Zjednoczonych. Obecnie najpopularniejsze systemy telewizji analogowej to NTSC i PAL. Mają one ze sobą wiele wspólnego. W zasadzie system PAL jest udoskonaloną modyfikacją systemu NTSC. Całkowity sygnał wizyjny jest na wielu etapach przetwarzania obrazu reprezentowany przez składniki komponentowe: luminancję Y i dwa sygnały różnicowe koloru R-Y oraz B-Y (sygnały chrominancji). Nie stosuje się przy transmisji i kompresji sygnałów RGB, gdyż każdy z nich jest sygnałem pełnej szerokości pasma. Sygnały różnicowe koloru mogą być natomiast ograniczone częstotliwościowo w stosunku do sygnału luminancji bez wpływu na jakość zrekonstruowanego obrazu barwnego. Próbkki chrominancji występują w strukturze linii (czyli w kierunku poziomym) dwukrotnie rzadziej niż elementy luminancji. Podobną zasadę można zastosować w kierunku pionowym, czyli umieszczać próbkki chrominancji na co drugiej linii. Takie ograniczenie rozdzielczości w pionie nie wpływa zasadniczo na jakość obrazu kolorowego, natomiast istotnie redukuje strumień informacji o obrazie. Schemat próbkowania określany jest wtedy jako 4:2:0.

System NTSC jest stosowany dla obrazów w formacie złożonym z 525 linii na ramkę, przy częstotliwości odświeżania 59,94Hz (jest to skutkiem stosowania w USA częstotliwości prądu przemiennego wynoszącej 60Hz) i 29,97 ramkach na sekundę. Odpowiedzialna za 2/3 informacji o barwie fala podnośna przenosi jednocześnie informacje o dwóch składowych danej barwy – czerwonej i zielonej. Powoduje to dużą podatność na zakłócenia fazowe.

Stosowany w Polsce standard telewizji kolorowej PAL bazuje na strukturze ramki obrazu zawierającej 625 linii i składającej się z dwóch pół półobrazów powtarzanych z częstotliwością 50Hz. W standardzie PAL stosuje się strukturę wybierania linii określaną jako wybieranie kolejnoliniowe nieparzyste. Linie należące do kolejnego półobrazu są wyświetlane na ekranie pomiędzy liniami poprzedniego. Każda pełna ramka obrazu pojawia się wobec tego 25 razy na sekundę.

4.2 STANDARD TELEWIZJI KOLOROWEJ HDTV

HDTV (ang. *High Definition TV*) to telewizja wysokiej rozdzielczości. W potocznym znaczeniu jest określeniem sygnału telewizyjnego o rozdzielczości większej niż standardowa (PAL lub NTSC). Pierwsze publiczne instalacje analogowej telewizji w wysokiej rozdzielczości zostały uruchomione w Japonii, gdzie wciąż cieszą się dużą popularnością, mimo równoległej transmisji w systemie cyfrowym. Podczas gdy w USA telewizja wysokiej rozdzielczości stawała się coraz popularniejsza, w Europie przez dłuższy czas nie była stosowana w publicznych przekazach. W końcu jednak w 2004 roku pojawiła się pierwsza stacja nadająca z europejskiego satelity Astra – euro1080. W Polsce pierwszym operatorem kablowym, który wprowadził usługę HDTV (w 2007 roku), były Multimedia Polska S.A.

HDTV oferuje rozdzielczości:

- 720p – 1280×720 pikseli



- 1080i/1080p – 1920×1080 pikseli
gdzie "i" (*interlaced*) oznacza obraz z przeplotem (na zmianę wyświetlane są linie parzyste i nieparzyste), po symbolu „i” czasami podawana jest liczba pól (półobrazów) na sekundę, np. 1080i60, natomiast p (*progressive scan*) oznacza obraz bez przeplotu. Po symbolu „p” podawana jest czasami liczba klatek (pełnych obrazów) na sekundę, np. 720p60.

Przed przesłaniem do użytkownika końcowego sygnał HDTV może być zakodowany na kilka sposobów, wśród których najczęściej stosuje się: MPEG-2, H.264/MPEG-4 AVC.

4.3 CYFROWA TELEWIZJA SYSTEMU DVB

DVB (ang. *Digital Video Broadcast*) to standard cyfrowej telewizji. Charakteryzuje się jakością obrazu i dźwięku porównywalną do zapisu DVD. Telewizja DVB umożliwia często interaktywny odbiór, np. włączenie napisów w różnych językach oraz przełączenia języka ścieżki audio. W standardzie DVB obraz i dźwięk są przesyłane w systemie MPEG-2, tak jak w DVD. Niedawno wprowadzono również kodowanie w MPEG-4 (H.264). Sygnał DVB może być przekazywany z nadajników naziemnych (DVB-T), satelity (DVB-S) i stacji telewizji kablowych (DVB-C). Podstawą tego systemu jest strumień transportowy (TS) zdefiniowany i opisany w normie międzynarodowej ISO/IEC 13818-1. TS składa się ze skompresowanych składowych wizji, fonii i danych oraz tablic (PSI) umożliwiających urządzeniu odbiorczemu odbiór wybranego programu telewizyjnego lub radiowego oraz danych. Standard DVB definiuje dodatkowe tablice (SI) umieszczone w strumieniu oraz parametry transmisji w zależności od typu kanału transmisyjnego.

4.4 POPRAWA JAKOŚCI OBRAZU

Najczęściej spotykane zniekształcenia wynikają z pojawienia się artefaktów procesu kompresji. Do zakłóceń zaliczamy między innymi: szумы, interferencje (przenikanie sygnałów luminancji i chrominancji), migotanie powierzchni i linii, zaburzenia synchronizacji. Eliminacja wymienionych zjawisk jest możliwa przy wykorzystaniu dwu- i trzymiwarowych filtrów cyfrowych, filtrów grzebieniowych, układów korekcji podstawy czasu i stosowaniu odpowiednich technik (100Hz, *Progressive Scan*). Poprawie jakości sprzyja też sztuczne podnośzenie rozdzielczości w oparciu o technikę nadpróbkowania i interpolacji wartości pikseli.

4.5 KINO DOMOWE

Kino domowe to zestaw urządzeń audio i wideo przeznaczony do oglądania filmów, który umożliwia symulowanie w warunkach domowych wrażenia jakiego doświadcza widz na sali kinowej. Zazwyczaj w skład zestawu wchodzi: odtwarzacz DVD, zestaw głośników wraz ze wzmacniaczem oraz telewizor. Współczesne systemy odtwarzania w domowych warunkach dźwięku przestrzennego wywodzą się z techniki filmowej. Rozwój tej techniki odbywał się na przestrzeni szeregu lat. Pierwsze formaty dźwięku wielokanałowego dla kina zostały opracowane na początku lat pięćdziesiątych.

Istnieje bardzo wiele standardów dźwięku przestrzennego. Jednym z pierwszych, analogowych systemów, był system **Dolby Stereo**, w którym na dwóch ścieżkach optycznych przy pomocy techniki matrycowej zakodowano cztery kanały foniczne. Oprócz tradycyjnych sygnałów lewego i prawego kanału mamy służący poprawnej lokalizacji dialogów kanał centralny oraz kanał *surround* przeznaczony do efektów specjalnych i tworzenia odpowiedniej atmosfery scenicznej. System Dolby Stereo został z powodzeniem wprowadzony do techniki filmowej i tysiące kin na całym świecie zostało wyposażone w stosowną aparaturę. System stworzył również podstawy do wprowadzenia dźwięku dookólnego do warunków domowych. Wprowadzony w 1987r. dekodery **Dolby Surround Pro Logic** umożliwiał wyodrębnienie z dwóch ścieżek czterech kanałów, podobnie jak miało to miejsce w aparaturze kinowej. Obecnie w systemie znanym jako Dolby Surround nadawane są nie tylko filmy kinowe ze ścieżką Dolby Stereo, ale także seriale telewizyjne oraz wydarzenia artystyczne i sportowe. Oba opisywane wyżej standardy są standardami analogowymi.

Pod koniec lat osiemdziesiątych Dolby Laboratories rozpoczęło prace nad wdrożeniem nowej cyfrowej technologii dźwięku dookólnego dla filmu 35 mm. w celu zapewnienia kompatybilności z dotychczasowym systemem analogowym, optyczną ścieżkę zawierającą cyfrową informację foniczną umieszczono wzdłuż perforacji taśmy filmowej. Nowy system nazwany **Dolby Digital** bazuje na sześciu dyskretnych (odseparowanych) kanałach dźwiękowych w konfiguracji oznaczonej symbolem 5.1. Konfiguracja ta składa się z pięciu pełnopasmowych kanałów: lewego, prawego, centralnego, lewego dookólnego i prawego dookólnego. Szósty dodatkowy kanał służy do przeniesienia efektów niskoczęstotliwościowych (*LFE – Low Frequency Effects*), które niekiedy w kinach są bardziej odczuwalne niż słyszalne (wrażenia towarzyszące trzęsieniu ziemi). Z tego



względu, że LFE wymaga około jednej dziesiątej pasma akustycznego odtwarzanego w pozostałych kanałach, jest on oznaczany jako kanał .1. Ograniczenie miejsca na taśmie filmowej na umieszczenie optycznego zapisu Dolby Digital wymusiło zastosowanie cyfrowej metody kompresji sygnału fonicznego AC-3 (*Audio Coding ver.3*). Dolby Digital debiutował w kinach w roku 1992 i do tej pory jest wiodącym cyfrowym formatem dźwięku przestrzennego.

Rozwój technologii cyfrowej w telewizji i technice wideo spowodował szybkie przyjęcie się formatu Dolby Digital w domowym sprzęcie audiowizualnym. W odróżnieniu od analogowego Dolby Surround Dolby Digital oferuje dwa niezależne kanały *surround* o jakości równoważnej trzem przednim kanałom. Dzięki skutecznej metodzie kompresji AC-3 format Dolby Digital został przyjęty jako standard dźwięku wielokanałowego w telewizji cyfrowej w USA. Większość płyt DVD również zawiera zapis ścieżki dźwiękowej w tym systemie.

Konkurencyjnym dla Dolby Digital systemem reprodukcji dźwięku wielokanałowego jest **DTS** (ang. *Digital Theater System*). Dzięki zastosowaniu mniejszego stopnia kompresji gwarantuje on potencjalnie lepszą jakość dźwięku. W odróżnieniu od kinowego systemu Dolby Digital ścieżki DTS z racji zwiększonej przepływności strumienia informacji nie są zapisywane na kliszy filmowej tylko na osobnym nośniku – płycie CD-ROM. Filmy zrealizowane w technice DTS zawierają w celu kompatybilności z aparaturą kinową starszego typu dwie analogowe ścieżki optyczne w formacie matrycowym DTS Surround, zgodnym z systemem Dolby Stereo. Do synchronizacji obrazu z taśmą filmową z dźwiękiem odtwarzanym z CD-ROM obok analogowych ścieżek optycznych umieszczono dodatkowy ślad z kodem czasowym. W systemie DTS również przyjęto konfigurację 5.1: trzy kanały przednie (lewy, prawy i centralny) przenoszą pełne pasmo akustyczne 20Hz – 20kHz, dwa kanały dookólne (lewy i prawy) działają w zakresie 80Hz – 20kHz, kanał LFE reprodukuje częstotliwości z pasma 20Hz – 80Hz. Rozważając strukturę zapisu DTS można stwierdzić, że jest to system pięciokanałowy, gdyż informacja dla głośnika niskotonowego jest zmiksowana w obu kanałach *surround*. Opcjonalnie w DTS można więc w prosty sposób wykorzystać pełnopasmowe głośniki dookólne bez stosowania dedykowanego subwoofera. System DTS również rozpowszechnił się w kinie domowym. Stosowany w DTS algorytm kompresji CAC (*Coherent Acoustic Coding*) daje przy 20-bitowej rozdzielczości i próbkowaniu 44,1kHz strumień o przepływności 1,4 Mbit/s, co oznacza stopień redukcji informacji nadmiarowej na poziomie 4:1. Mały stopień kompresji sprzyja podniesieniu jakości dźwięku, z drugiej strony uniemożliwia zastosowanie tego systemu w standardach transmisji i zapisu fonii wymagających dużo mniejszych przepływności.

Dążenie do zwiększenia realizmu odtwarzanych dźwięków dochodzących z otoczenia słuchacza spowodowało pojawienie się systemów o większej niż sześć liczbie kanałów. W technice **Dolby EX 6.1** wprowadzony został dodatkowy kanał centralny *surround* sterujący głośnikiem umieszczonym za głową słuchacza. Sygnał tego kanału nie jest przesyłany w sposób dyskretny, lecz jest zakodowany matrycowo wraz z informacją z kanałów dookólnych lewego i prawego. Subiektywnie lepsze efekty daje zaproponowany przez **Dolby Laboratories format 7.1**, który operuje czterema głośnikami dookólnymi: bocznym lewym, bocznym prawym, tylnym lewym i tylnym prawym. Taka konfiguracja w porównaniu z 6.1 pozwala uniknąć występującego czasami zjawiska pozornej kumulacji obrazu dźwiękowego w środku pomieszczenia odsłuchowego. Dzięki kodowaniu matrycowemu nagrania w formacie 6.1 i 7.1 mogą być odtwarzane przez dekodery Dolby Digital 5.1, tak jak ścieżka dźwiękowa 5.1 może być reprodukowana w systemach Dolby EX z użyciem większej liczby głośników. Również firma DTS rozpoczęła prace nad nowymi formatami. W systemie DTS 6.1 Discrete wprowadzono dodatkowy niezależny centralny kanał tylny. Ścieżka foniczna w obecnie stosowanym formacie DTS-ES 6.1 zawiera także informacje tego kanału zakodowane matrycowo w dwóch pozostałych kanałach dookólnych. Użytkownik systemu 6.1 w zależności od posiadanego dekodera może wyodrębnić sygnał tylnego kanału centralnego z dyskretnej ścieżki lub poprzez dekodowanie matrycowe z sygnałów *surround*. Projektanci z firmy DTS w dążeniu do dalszej poprawy jakości dźwięku wprowadzili nowy format DTS 96/24. Zgodnie z przyjętą przez firmę koncepcją odtwarzania dźwięku towarzyszącemu obrazowi z jakością lepszą niż CD, w miejsce konwencjonalnego próbkowania sygnału audio z częstotliwością 44.1kHz ze słowem 16-bitowym, zaproponowano próbkowanie 96kHz z rozdzielczością 24 bitów. Dłuższe słowo bitowe zwiększa zakres dynamiczny nagrania, natomiast większa częstotliwość próbkowania umożliwia przetwarzanie sygnałów o szerszym paśmie i ułatwia pracę filtrom antyaliasingowym na wyjściach przetworników cyfrowo-analogowych. System DTS 96/24 jest wstecznie kompatybilny z wcześniejszymi dekodernami o częstotliwości taktowania 48kHz. Sygnał w formacie 96/24 jest dostępny na wyjściach cyfrowych odtwarzaczy i może być przetwarzany przez wysokiej klasy zewnętrzne przetworniki C/A.



4.6 AKUSTYKA POMIESZCZENIA

W celu uzyskania zadowalających efektów dźwiękowych powinniśmy wybrać pomieszczenie o odpowiednim kształcie i wielkości. Fale dźwiękowe emitowane przez głośniki rozchodzą się w wielu kierunkach. Duża część dźwięków dociera do słuchacza po kilkukrotnym odbiciu się od ścian, sufitu czy podłogi. Zalecane jest więc również odpowiednie wyttumienie określonych części pokoju przeznaczonego do odsłuchu kina domowego.

Odbicia dźwięku nie są tylko i wyłącznie efektem negatywnym. Całkowite wyciszenie pomieszczenia sprawi, że słuchany dźwięk będzie brzmiał nienaturalnie. Natomiast w przypadku niewyttumionych „gołych” ścian i dodatkowo złego kształtu pomieszczenia, pojawiają się rezonanse zaburzając charakterystykę częstotliwościową. Naturalnie rolę tłumiącą spełnia umeblowanie pomieszczenia (w przypadku gdy korzystamy z miejsca nie przeznaczonego tylko i wyłącznie do odsłuchu kina domowego). Tony niskie są skutecznie pochłaniane przez drewniane konstrukcje które absorbują drgania. Tony średnie i wysokie tłumione są przez materiały miękkie (np. dywany).

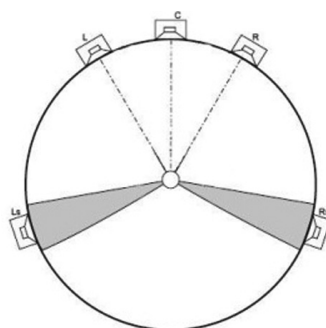
Jednym z najważniejszych pojęć determinujących akustykę pomieszczenia jest **rezonans**. W pomieszczeniu zamkniętym fale dźwiękowe odbijają się od ścian. W przypadku gdy fale nachodzą na siebie niektóre częstotliwości mogą zostać wzmocnione a inne stłumione. Przy pewnych częstotliwościach nakładające się na siebie fale o przeciwnych kierunkach tworzą fale stojące. Najbardziej podatne na to zjawisko są pomieszczenia o równoległych płaskich ścianach. Jeśli tylko jest to możliwe ściany pokoju do odsłuchu kina domowego powinny być nierównoległe o nieregularnym kształcie, a sufit lekko nachylony. W przypadku braku możliwości zrealizowania tych założeń zaleca się aby narożniki obiektu były zaokrąglone. Bardzo istotne jest dobranie odpowiednich proporcji pomieszczenia (wysokość, szerokość, długość). Zależy od tego czy rezonanse są równomiernie rozłożone, czy też nachodzą na siebie. Optymalne proporcje pomieszczenia do odsłuchu dźwięków przedstawił amerykański akustyk L.W. Sepemeyer 1 :1,14 : 1,39; 1: 1,28 : 1,54 oraz 1 : 1,60 : 2,33 przy czym 1 to wysokość pomieszczenia, a dwie kolejne wartości to odpowiednio szerokość i długość.

Szukając informacji na temat idealnych kształtów pokoju odsłuchowego/kinowego możemy trafić na wiele propozycji. Poczynając od ogólnego stwierdzenia o „Braku równoległych ścian” kończąc na przykładach kształtów, takich jak: sześciokąty, trapezy czy formy owalne. Takie kształty miały by teoretycznie niwelować fale stojące oraz rezonanse wynikające z odbijania się fal dźwiękowych od równoległych ścian, lecz wielu ekspertów twierdzi, że nie rozwiązują one problemu, a wręcz wprowadzają nowe.

Trzeba również zwrócić uwagę na wielkość pokoju. Zestawy głośnikowe, a dokładniej ich moc, powinny być dobrane zgodnie z wielkością pokoju by dobrze wypełnić powierzchnie dźwiękiem. W małych pokojach trzeba pogodzić się z gorszą jakością dźwięku, a w szczególności basu i stereofonii. Częściowym wyjściem jest odpowiednie ustawienie zestawu głośnikowego, lecz na małym metrażu zazwyczaj będzie ono kolidowało z rozmieszczeniem umeblowania oraz innych sprzętów, a końcowy efekt będzie nadal gorszy niż na większych powierzchniach. Za optymalną powierzchnię można uznać 25-35m².

Odpowiednie rozmieszczenie kolumn jest najtańszą, a za razem najbardziej wpływającą na efekt końcowy czynnością przy projektowaniu pokoju dla kina domowego. W większości przypadków nie kosztuje to zupełnie nic, a w pozostałych przypadkach cena nie jest wysoka. Dlatego właśnie warto poświęcić czas, a czasami powierzchnie użytkową dla uzyskania optymalnych efektów.

Istnieje wiele metod ustawiania głośników. Sposób rozstawiania głośników w popularnym standardzie 5.1 jest zilustrowane na rys. 42.



C – kolumna centralna
L, R – kolumny przednie
L_s, R_s – kolumny efektowe

Rysunek 42.
ozmieszczenie głośników w standardzie 5.1



Wszystkie trzy główne zestawy głośnikowe (L – lewy, C – centralny, R – prawy) powinny znajdować się w równej odległości od słuchacza. W przypadku gdy fizycznie niemożliwe jest ich odpowiednie rozstawienie, niezbędne jest użycie systemów regulujących opóźnienia w poszczególnych kanałach. Systemy takie są dostępne w amplitunerach dobrej klasy. Każde 0,3 metra odległości to odpowiednik 1ms czasu opóźnienia. Dodatkowo trzy główne kolumny powinny znajdować się na tej samej wysokości tak, aby głośniki wysokotonowe znalazły się na poziomie uszu słuchacza. Ze względu na użycie w większości przypadków wyświetlacza ekranowego, takie ustawienie głośnika centralnego jest zazwyczaj niemożliwe. Należy wtedy umieścić głośnik centralny nad ekranem równocześnie nachylając go ku podłodze tak, aby fale dźwiękowe rozchodziły się dokładnie w kierunku osoby słuchającej. Należy również przemyśleć położenie zestawów głośników w stosunku do ścian bocznych i ściany tylnej. Odsunięcie od ściany tylnej często wymuszane jest przez konstrukcję zestawu (głośniki bipolarne lub z otworem bassreflex umieszczonym na tylnej ścianie głośnika). Odsunięcie od ściany tylnej jak i bocznej wpływa w znaczący sposób na charakter dźwięków. Użytkownik powinien dokonać tego wyboru według własnych upodobań. Nie powinno się jednak stawiać zespołów głośnikowych bliżej niż 0,5 metra od ścian bocznych i 1 metr od ściany tylnej. Należy również pamiętać aby odległość od słuchacza do ściany za nim nie była równa odległości od głównych zestawów głośnikowych do ściany za nimi, co powodowałoby zrównanie długości drogi fali odbitej od obu ścian. Jeśli chodzi o rozstawienie głośników przednich L i R, to odległość między nimi nie powinna przekraczać odległości od słuchacza. W przypadku gdy zestaw ma głośnik centralny, większy rozstaw głośników powodować będzie mniejsze straty niż ustawienie ich zbyt blisko siebie. Zbyt blisko ustawione zestawy głośnikowe zawężają scenę dźwiękową.

Głośniki surround montujemy po bokach osoby słuchającej. Kąt pomiędzy osią słuchacz – głośnik centralny, a osią słuchacz – głośnik surround powinien zawierać się w przedziale 90-110 stopni. Wysokość montażu głośników bocznych zależy od wielkości pomieszczenia przeznaczonego do odsłuchu kina domowego. Głośniki te powinny być umieszczone nad głową słuchacza i skierowane na niego tak aby kąt ich odchylenia od poziomu zawierał się w przedziale 15-20 stopni. W celu uzyskania najlepszych efektów należy użyć głośników bocznych dipolowych. Głośniki te kierują fale dźwiękowe do przodu i do tyłu pomieszczenia a nie bezpośrednio na słuchacza. Dźwięk trafia do niego dopiero po odbiciu się od ściany przedniej i tylnej.

Tylne głośniki surround powinny zostać zamontowane na tej samej wysokości co boczne. Umieszczone powinny być obok siebie prosto za osobą odsłuchującą efekty dźwiękowe.

Głośniki średnio i wysoko tonowe powinny być widoczne przez słuchacza. Na drodze pomiędzy miejscem odsłuchowym a zestawami głośnikowymi nie powinny znajdować się przeszkody. Zasada ta nie dotyczy jednak głośnika niskotonowego (*subwoofer*). Wiąże się z nim zjawisko wszechkierunkowości. Oznacza to, że dźwięk z tego głośnika rozchodzi się w każdym kierunku, a przeszkody stojące na drodze do słuchacza nie mają dużego wpływu na jakość dźwięku. Subwoofer wywołuje jednak różne zjawiska rezonansowe. Korzystne jest zastosowanie dwóch tego typu głośników stojących naprzeciwko. Zapewnia to wyrównanie rezonansu i zwiększenie płynności basu. W przypadku posiadania tylko jednego głośnika niskotonowego zaleca się umieszczenie go przy ścianie przedniej w odległości od ściany bocznej równej 1/3 szerokości pomieszczenia.

Każde pomieszczenie przeznaczone do odsłuchu kina domowego jest odrębnym przypadkiem. Różna wielkość, kształt, charakterystyka akustyczna uniemożliwiają ustalenie „złotej receptury” rozmieszczania głośników. Zastosowanie powyższych wskazówek na pewno zwiększy wydajność zestawu. Aby uzyskać idealną jakość dźwięku należy jednak dopracować rozstawienie zestawów głośników metodą eksperymentalną, wiele bowiem zależy od gustu osoby konfigurującej swoje własne kino domowe.

Zadanie 7. Pasma przenoszenia głośników, krzywa słyszalności

Do wykonania zadania jest potrzebny:

- komputer z zainstalowaną kartą dźwiękową,
- zestaw głośników komputerowych,
- zestaw kina domowego.

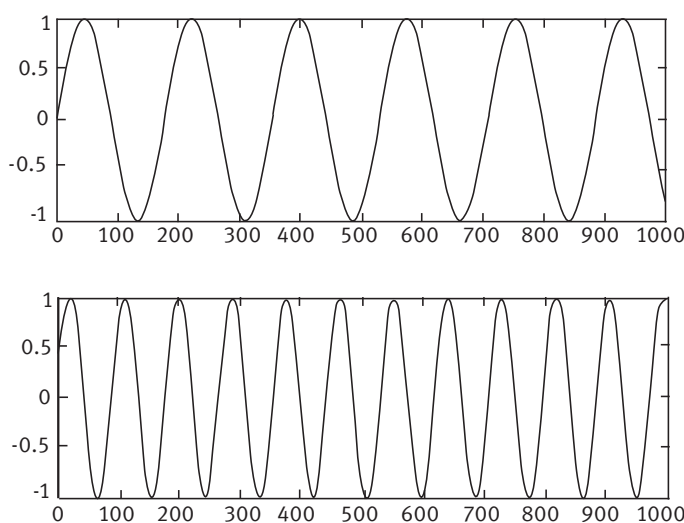
W tym zadaniu przeprowadzona będzie demonstracja pasma przenoszenia głośniczków komputerowych i jednocześnie krzywej słyszalności (w zakresie pasma przenoszenia głośników). Ta sama demonstracja przeprowadzona będzie następnie dla zestawu kina domowego. Polega ona na odtworzeniu sygnałów sinusoidalnych o tej samej amplitudzie i różnych częstotliwościach (rys. 43). Należy porównać pasma przenoszenia różnych zestawów sprzętowych.



Poszczególne prezentowane częstotliwości to: 40Hz, 60Hz, 100Hz, 500Hz, 2000Hz, 3000Hz, 8000Hz, 10000Hz, 16000Hz. Sygnały mają takie same amplitudy, więc wydawałoby się, że będą odbierane z taką samą głośnością. Jednak skrajne częstotliwości są bardzo słabo słyszalne. Sygnał o częstotliwości 3000Hz słychać najlepiej. Słuchając prezentowanych dźwięków można się łatwo zorientować jaki zakres częstotliwościowy sygnałów jest przenoszony przez układ komputer, karta dźwiękowa, głośniczki.

W drugiej części będziemy słuchać sygnałów sinusoidalnych o różnych amplitudach, np. 1, 2, 10, 100, ale o tej samej częstotliwości. Należy zaobserwować, czy sygnał o amplitudzie dwa razy większej będzie odbierany jako dwa razy głośniejszy. Dlaczego tak nie jest? Jak należy zwiększyć amplitudę sygnału dźwiękowego, żeby usłyszeć go jako dwa razy głośniejszy? Co to jest dB i skala logarytmiczna?

Po demonstracji słuchacze powinni zapoznać się z możliwościami konfiguracji zestawu kina domowego. Należy wiedzieć jakie elementy wchodzi w skład zestawu i jak je połączyć. Następnie należy zastanowić się, czy pomieszczenie odsłuchowe, którym dysponujemy, jest korzystne pod względem akustyki. Jak najlepiej ustawić kolumny? Jakie ustawienie kolumn powoduje, że dźwięk jest nie do zaakceptowania? Eksperymentalnie należy dobrać najlepsze rozstawienie kolumn głośnikowych.



Rysunek 43.

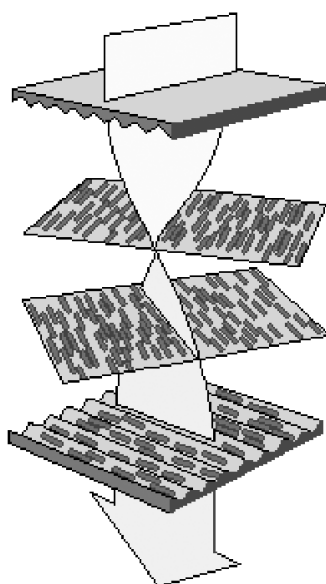
Dwa sygnały sinusoidalne o tych samych amplitudach przy czym częstotliwość pierwszego sygnału jest dwa razy mniejsza niż drugiego

4.7 WYŚWIETLACZE LCD

Ciekłe kryształy zostały wynalezione w XIX wieku przez austriackiego botanika Friedricha Reinitzera, a termin rozpropagował niemiecki fizyk Otto Lehmann. Ciekłe kryształy to substancje prawie przezroczyste, mające właściwości charakteryzujące zarówno ciała stałe, jak i ciecze. Światło przechodzące przez ciekłe kryształy podąża za ułożeniem tworzących je molekuł. W 1960 roku odkryto, że pobudzenie napięciem elektrycznym ciekłych kryształów zmienia położenie tych molekuł, a co za tym idzie – sposób przenikania przez nie światła. W roku 1969 James Fergason odkrył efekt skręconego **nematyka** (ang. *twisted nematic* – *TN*). Było to odkrycie o fundamentalnym znaczeniu, ponieważ wyświetlacze LCD wykorzystują właśnie to zjawisko.

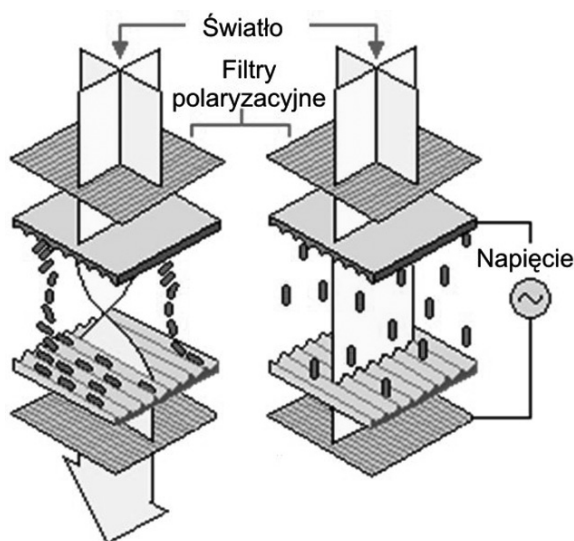
Panele LCD świecą dzięki zastosowaniu specjalnych lamp z tzw. zimną katodą. Charakteryzują się one bardzo dużą wydajnością przy jednoczesnym niewielkim zużyciu energii. Użycie filtra polaryzującego światło powoduje polaryzację przechodzącej przez niego wiązki światła. Polaryzacja światła zależy od orientacji wektora jego pola elektrycznego. W uproszczeniu światło to fala elektromagnetyczna. Wektory pola elektrycznego i magnetycznego są prostopadłe do kierunku fali ruchu. Lampa emituje nie spolaryzowane światło, więc pole elektryczne może poruszać się w dowolnym kierunku prostopadłym do osi propagacji światła. Gdy światło przechodzi przez polaryzator, to wychodząc po drugiej stronie ma wektor pola elektrycznego skierowany w znanym kierunku (np. pionowym). Gdy światło przechodzi przez drugi polaryzator, prostopadły do pierwszego (w tym przypadku poziomy), to nie może przez niego przejść. Ale jeżeli między dwoma polaryzatorami umieści się ciekły kryształ, to zmienia on polaryzację światła na pasującą do drugiego polaryzatora i wtedy światło jest przepuszczane przez cały układ.





Rysunek 44.
Zasada działania wyświetlacza LCD w technologii TN

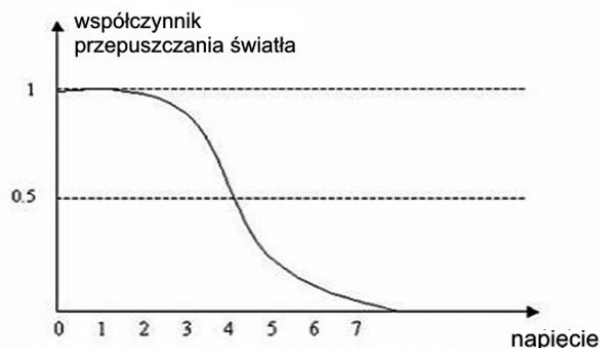
W technologii TN ekran LCD składa się z dwóch warstw ciekłych kryształów umieszczonych pomiędzy dwiema odpowiednio wyprofilowanymi powierzchniami (rys. 44), z których jedna jest ustawiona najczęściej pod kątem 90 stopni wobec drugiej (stąd w nazwie skręcenie – *twisted*). Molekuły znajdujące się między nimi muszą się przemieścić o 90 stopni, podobnie jak światło podążające za ich położeniem. Wystarczy jednak przyłożyć do ciekłych kryształów napięcie elektryczne, a molekuły zaczną się przemieszczać pionowo, powodując przejście światła bez zmiany położenia o 90 stopni (rys. 45). Technologia TN jest najczęściej stosowana w niedrogich modelach monitorów komputerowych o niewielkich przekątnych obrazu – 15, 17 cali. Matryce tego typu nie najlepiej reprodukuje barwy, mają jednak również mocne strony. Gwarantują na przykład krótki czas reakcji – w najnowszych modelach 4ms. Największą wadą matryc TN jest stosunkowo wąski kąt widzenia, 120-140 stopni w obu płaszczyznach. Panele tego typu nie nadają się do zastosowań profesjonalnych (pracy z grafiką), jednak z racji niskiej ceny i wspomnianego krótkiego czasu reakcji dobrze sprawdzą się podczas oglądania dynamicznych materiałów wideo.



Rysunek 45.
Napięcie powoduje zmianę położenia molekuł ciekłego kryształu



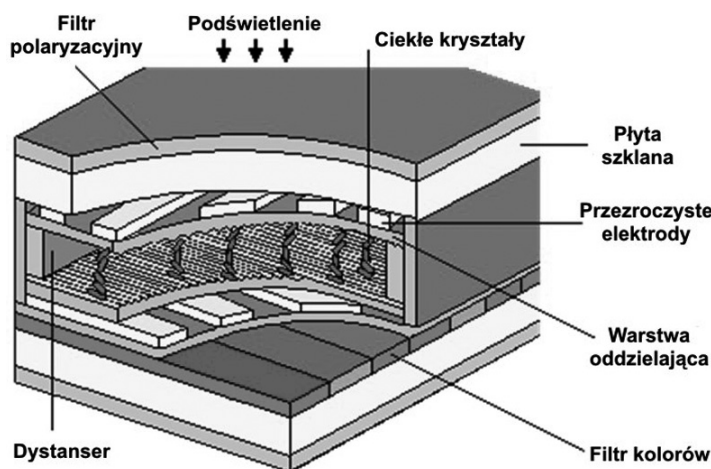
Jedną z wad paneli ciekłokrystalicznych jest gorsze odwzorowanie barw w stosunku do monitorów CRT. Ma to szczególnie znaczenie przy profesjonalnej obróbce zdjęć, zwłaszcza, gdy wyniki naszej pracy mają trafić do drukarni. Normalny obserwator odróżnia od 300.000 do 1.000.000 barw natomiast przy pomocy techniki cyfrowej można na monitorze komputera przedstawić obraz w modelu RGB, mający dokładność 24 bitów (3 kanały x 8 bitów). Każdy kanał barwy R, G, B jest opisany liczbą 8 bitową, dzięki czemu można uzyskać w każdym kanale 256 poziomów jasności. Kombinacja 256 poziomów koloru czerwonego, zielonego i niebieskiego definiuje przeszło 16 milionów kolorów. Różnicując napięcie na końcówkach ciekłego kryształu można modulować stopień zamknięcia przełącznika, aby uzyskać stany pośrednie (rys. 46). Niestety w technologii TN można w ten sposób uzyskać jedynie rozdzielczość ok. 6 bitów (8 bitów w technologii CRT). Pozostałe, brakujące barwy są zazwyczaj uzyskiwane na zasadzie interpolacji z lepszym lub gorszym skutkiem.



Rysunek 46.

óżnicując napięcie na końcówkach ciekłego kryształu można modulować stopień zamknięcia przełącznika, aby uzyskać stany pośrednie

W **matrycach pasywnych** (rys. 47) ciekłe kryształy są adresowane poprzez ładunki lokalne, przy czym nic nie powstrzymuje ładunków elektrycznych przed rozptyływaniem się na boki i wpływaniem na położenie molekuł kryształów sąsiednich. Adresowanie, czyli określenie piksela, który ma być w danej chwili wysterowany, realizowane jest przy użyciu dwóch krzyżujących się elektrod. Elektroda przednia jest wspólna dla całej kolumny i przewodzi prąd, natomiast tylna elektroda, wspólna dla całego rzędu, służy jako uziemienie. Czas reakcji matrycy musi być bardzo długi, nawet kilkaset milisekund, gdyż ciekły kryształ musi zachować orientację molekuł do następnego zaadresowania. Nic nie podtrzymuje orientacji molekuł stąd powoli wracają one do położenia pierwotnego. Matryce pasywne charakteryzują się rozmytym obrazem oraz smugami i cieniami ciągnącymi się za obiektami w ruchu.

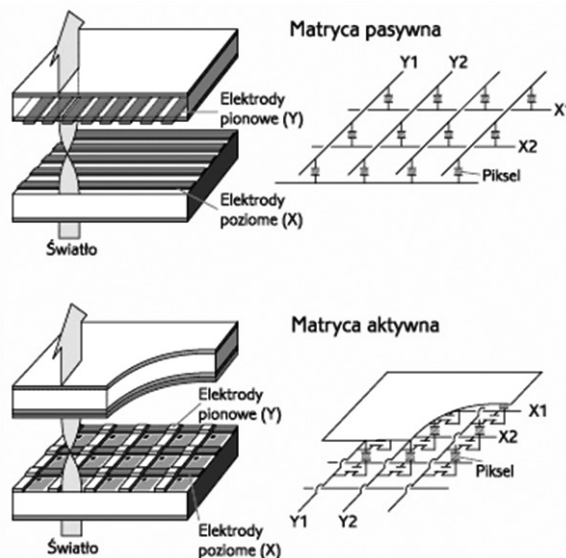


Rysunek 47.

Budowa matrycy pasywnej



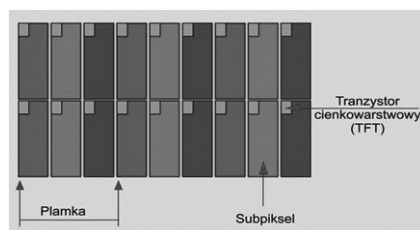
Matryce aktywne mają podobną budowę do matryc pasywnych. Podstawową różnicę stanowi warstwa tranzystorów cienkowarstwowych (ang. *thin film transistor, TFT*). Poprzez te tranzystory sterowane są kondensatory, które gromadzą i utrzymują ładunki elektryczne, zapobiegając ich rozptyłowaniu się na inne piksele (rys. 48).



Rysunek 48.
Budowa matrycy aktywnej

Tranzystor powiązany jest tylko z jedną komórką ciekłego kryształu, dzięki czemu nie ma smużenia ani rozmycia obrazu. Jeśli do elektrod przyłożymy napięcie, to spowodujemy, że cząsteczki ciekłych kryształów zmienią położenie i zostaną skrzyżowane. Zapisany ładunek pozostaje na kondensatorze i dzięki temu na końcówkach kryształów nadal jest napięcie, nawet gdy linia adresuje inny piksel. Oznacza to, że nie powróci on do stanu początkowego, co miało miejsce w przypadku matryc pasywnych. Czas zapisu ładunku do kondensatora jest o wiele krótszy niż czas obrotu kryształu, co oznacza, że dane mogą być zapisane, a kolejny piksel zaadresowany natychmiast, bez opóźnień. Obecnie stosuje się praktycznie wyłącznie matryce aktywne.

Specjalne filtry nadają kolor poszczególnym subpikselom ułożonym na przedniej warstwie szkła (rys. 49). Trzy subpiksele, każdy w kolorze czerwonym, zielonym oraz niebieskim, formują piksel. Różne kombinacje kolorystyczne subpikseli dają obraz oraz kolor na ekranie.



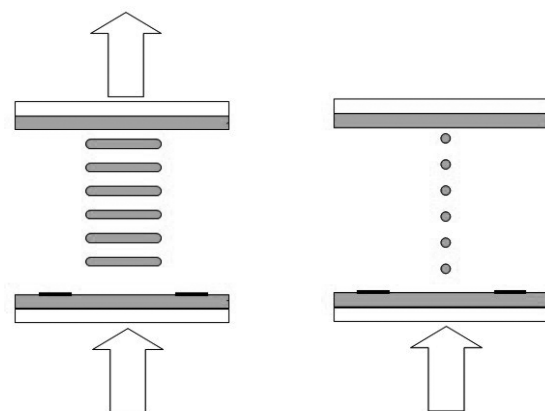
Rysunek 49.
Sposób nakładania filtrów barwnych

Oprócz omawianej wcześniej technologii TN wytwarzania wyświetlaczy LCD istnieją jeszcze dwie alternatywy: IPS, MVA.

Technologia **IPS, S-IPS** (ang. *In-Plane Switching*) została zaprojektowana w 1995 r. przez firmę Hitachi. Tym, co różni wyświetlacze IPS od wykonanych w technologii TN jest równoległe do powierzchni ułożenie molekuł ciekłych kryształów. Przy użyciu technologii IPS osiągnięty jest doskonały kąt widzenia, aż do 170 stopni,

jaki znamy z normalnych monitorów (CRT). Jednakże jest też minus: z powodu równoległego ułożenia ciekłych kryształów, elektrody nie mogą być umieszczone na obydwu szklanych powierzchniach jak w wyświetlaczu TN. Zamiast tego, muszą być zainstalowane w kształcie grzebienia na dolnej niższej powierzchni szklanej. Prowadzi to ostatecznie do redukcji kontrastu i dlatego wymagane jest silniejsze tylne podświetlenie dla podniesienia jasności obrazu. Obecnie S-IPS ma zalety matryc VA (ładne kolory, szerokie kąty widzenia) oraz TN (szybkość działania).

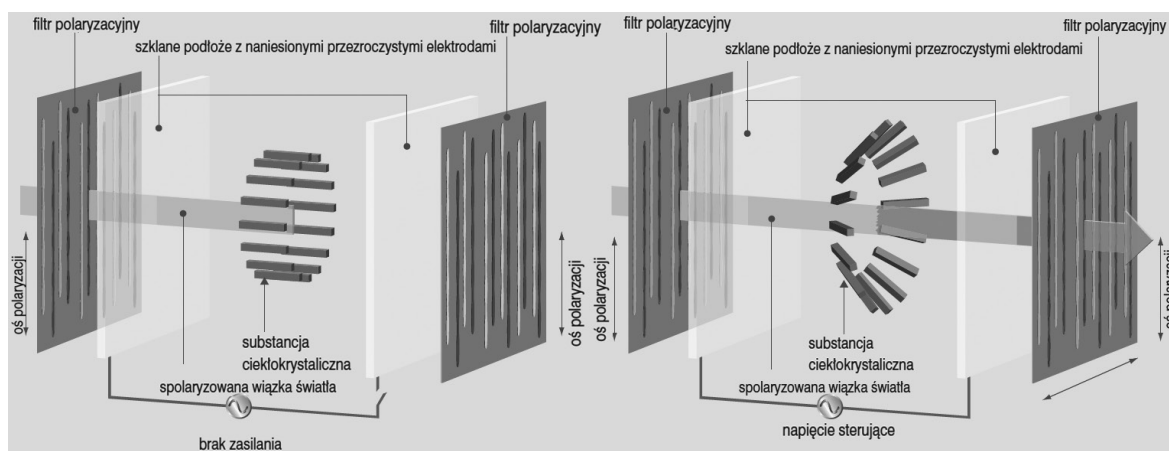
Na rys. 50 z lewej strony przedstawiono stan, w którym piksel wyświetlacza jest jasny (stan włączenia). Po przyłożeniu napięcia molekuly kryształu zaczynają się obracać o 90° pozostając w pozycji równoległej do siebie i do powierzchni wyświetlacza. Takie ułożenie molekuł ciekłego kryształu powoduje, że światło jest blokowane.



Rysunek 50.

Sposób ułożenia molekuł ciekłego kryształu w technologii IPS

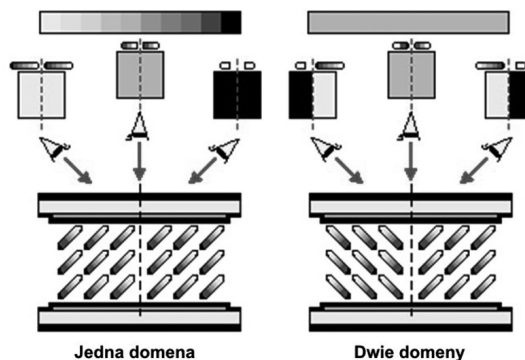
MVA, PVA, WVA to matryce uznawane za najlepiej odwzorowujące barwy. W popularnych matrycach TN ułożenie ciekłych kryształów będących w stanie pobudzenia jest nieuporządkowane, natomiast w wypadku tych technologii – wielokierunkowe, dzięki czemu obraz jest spójny, o mniej widocznej strukturze pikseli (rys. 51). Technologię PVA opracowała firma Samsung, natomiast MVA i WVA opatentowali inni znaczący producenci matryc LCD, Fujitsu i CMO. Zasada działania trzech technologii VA jest praktycznie taka sama, a różnice w nazewnictwie są spowodowane kwestiami praw patentowych. Panele tego typu mają bardzo szerokie kąty widzenia, minimum 170 stopni w obu płaszczyznach. Niestety mają też jedną poważną wadę – długi czas reakcji. Najszybsze modele MVA mają czas reakcji rzędu 25 ms, co sprawia, że podczas wyświetlania szybko przemieszczających się po ekranie obiektów można zauważyć smużenie.



Rysunek 51.

Budowa wyświetlaczy LCD wykonanych w technologii MVA

Na rys. 52 przedstawiono zasadę działania matrycy MVA. W tym przypadku zastosowane są tylko dwie domeny, w praktyce muszą być co najmniej cztery. Światło dochodzi do obserwatora w momencie kiedy molekuly ciekłego kryształu są ułożone prostopadle do kierunku obserwacji, a zatrzymywane jest w przypadku równoległego ułożenia molekuł. W położeniach pośrednich przepuszczana jest tylko część światła. W przypadku jednej domeny (rys. 52 po lewej) obserwator przechodząc z lewej strony na prawą widzi stopniowy zanik świecenia piksela. W przypadku zastosowania dwóch domen (rys. 52 po prawej) z lewej strony jedna domena przepuszcza światło druga natomiast je blokuje, z prawej strony jest odwrotnie. Gdy obserwator patrzy na wprost obie domeny przepuszczają tylko część światła. Ponieważ piksele są bardzo małe, oko i mózg ludzki uśredniają czarno-białe pola z lewej i prawej strony do koloru szarego. Obserwator w tym przypadku widzi jednolity szary kolor w bardzo szerokim zakresie kątów widzenia.



Rysunek 52.
Zasada działania matrycy MVA

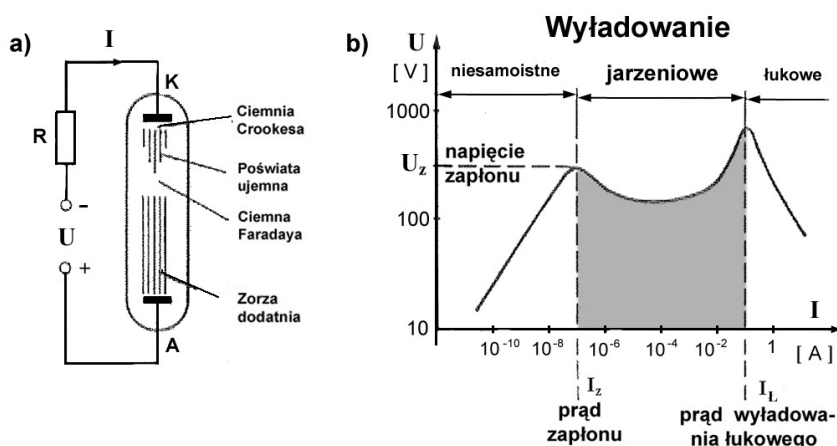
Podstawowe parametry charakteryzujące monitory ciekłokrystaliczne:

- **Rozdzielczość** – monitor LCD w przeciwieństwie do modeli CRT pracuje z maksymalną jakością tylko w rozdzielczości rzeczywistej (natywnej). Oczywiście prezentacja obrazu z inną rozdzielczością jest możliwa, jednak wtedy mamy do wyboru dwa sposoby oglądania obrazu – wyświetlany na fragmencie matrycy odpowiadającej danej rozdzielczości (np. 640x480 na panelu o rzeczywistej rozdzielczości 1024x768) lub prezentowany na całej powierzchni ekranu przy użyciu algorytmów interpolowania.
- **Częstotliwość odświeżania obrazu** w monitorach LCD – bezwładność monitorów ciekłokrystalicznych jest znacznie większa niż monitorów CRT. Każdy piksel matrycy LCD jest aktywowany oddzielnie i znajduje się w stanie włączonym albo wyłączonym. Obraz na monitorze LCD nie migocze, nie ma więc potrzeby niwelowania efektu migotania przez zwiększanie częstotliwości odświeżania. Częstotliwość odświeżania monitorów LCD dobiera się tak, aby zapewnić płynne zmiany obrazu przy animacji. Do tego celu w zupełności wystarczy odświeżanie z częstotliwością 60Hz i z taką częstotliwością pracuje większość monitorów LCD.
- **Czas reakcji** – producenci prezentują wyniki czasu reakcji piksela przy przejściu trzech subpikseli (zielony, czerwony, niebieski) od koloru czarnego do białego i odwrotnie. Suma czasów zapalania i gaszenia piksela składa się na czas końcowy, podawany w milisekundach (ms). Warto jednak zauważyć, że takie przedstawienie sprawy nie w pełni informuje, jak monitor będzie funkcjonował w praktyce. Rzeczywisty czas reakcji będzie taki, jak czas przejścia najwolniejszego z subpikseli, których kombinacje tworzą poszczególne kolory.
- **Kąty widzenia** – producenci paneli mierzą wielkość kątów widzenia poprzez utratę jakości obrazu. Moment, w którym następuje zbyt duża utrata jasności i kontrastu obrazu w porównaniu do wyjściowej, staje się kątem granicznym. Do mierzenia jakości kontrastu w monitorach LCD używa się współczynnika CR (ang. *Contrast Ratio*). W przeszłości kąty widzenia w monitorach LCD były bardzo małe, uniemożliwiając pracę więcej niż jednej osobie na monitorze. Po wprowadzeniu do masowej produkcji paneli w technologiach IPS oraz MVA, problem ten częściowo zniknął.

4.8 EKRANY PLAZMOWE

Ekrany plazmowe należą do grupy przetworników z wyświetlaniem aktywnym, wykorzystujących do wyświetlania zjawisko wyładowania jarzeniowego w plazmie. Istota tego zjawiska polega na emisji światła przez

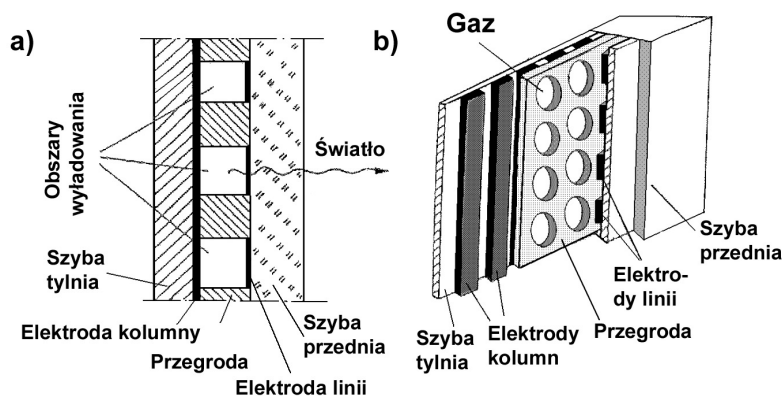
zjonizowany gaz o małym ciśnieniu (rzędu 1hPa) wskutek przepływu przez gaz prądu elektrycznego. Świecenie gazu jest wywołane zderzeniami jonów, początkowo samoistnie występujących, przyspieszanych w polu elektrycznym występującym pomiędzy dwiema elektrodami wyładowczymi spolaryzowanymi napięciem U . Dla małych napięć U (rys. 53) świecenie jest słabe, prawie niedostrzegalne. W miarę wzrostu napięcia U liczba jonów rośnie, co powoduje, że świecenie jest coraz intensywniejsze. Po przekroczeniu pewnego napięcia progowego U_z , zwanego napięciem zapłonu, cały gaz w obszarze pomiędzy elektrodami wyładowczymi jest zjonizowany – czyli tworzy tzw. plazmę (stąd nazwa przetwornika) – i świeci równomiernym światłem. Barwa wyładowania jarzeniowego zależy od rodzaju zastosowanego gazu.



Rysunek 53.

Zjawisko wyładowania jarzeniowego w plazmie

Na rys. 54 jest przedstawiona zasada konstrukcji najprostszego plazmowego wyświetlacza obrazów wykorzystującego opisane wyżej zjawisko fizyczne. Dwa zestawy elektrod ułożone prostopadle względem siebie są naniesione na wewnętrzne powierzchnie szklanych płyt, tworzących obudowę przetwornika, pomiędzy którymi znajduje się rozrzedzony gaz. Adresowanie ekranu polega na dołączeniu do adresujących elektrod, poprzez rezystor ograniczający prąd wyładowania, napięcia stałego większego od napięcia zapłonu. Wówczas w obszarze przestrzennego skrzyżowania adresowanej linii i adresowanej kolumny powstają warunki dla „zapłonu” i z węzła zostanie wyemitowane światło. Dla zapobieżenia rozprzestrzeniania się wyładowania do sąsiednich węzłów są one od siebie odseparowane przegrodą. Opisane wyżej rozwiązanie, od sposobu sterowania jest nazywane **stałoprądowym ekranem** (wyświetlaczem) plazmowym (DC PDP). Nie jest to rozwiązanie dogodne. Istotnymi jego wadami są: bezpośredni styk elektrod sterujących ze świecącym gazem, co powoduje ich stopniowe zniszczenie, konieczność stosowania rezystorów ograniczających prąd wyładowania, a także poważne trudności z wykonaniem przegrody.

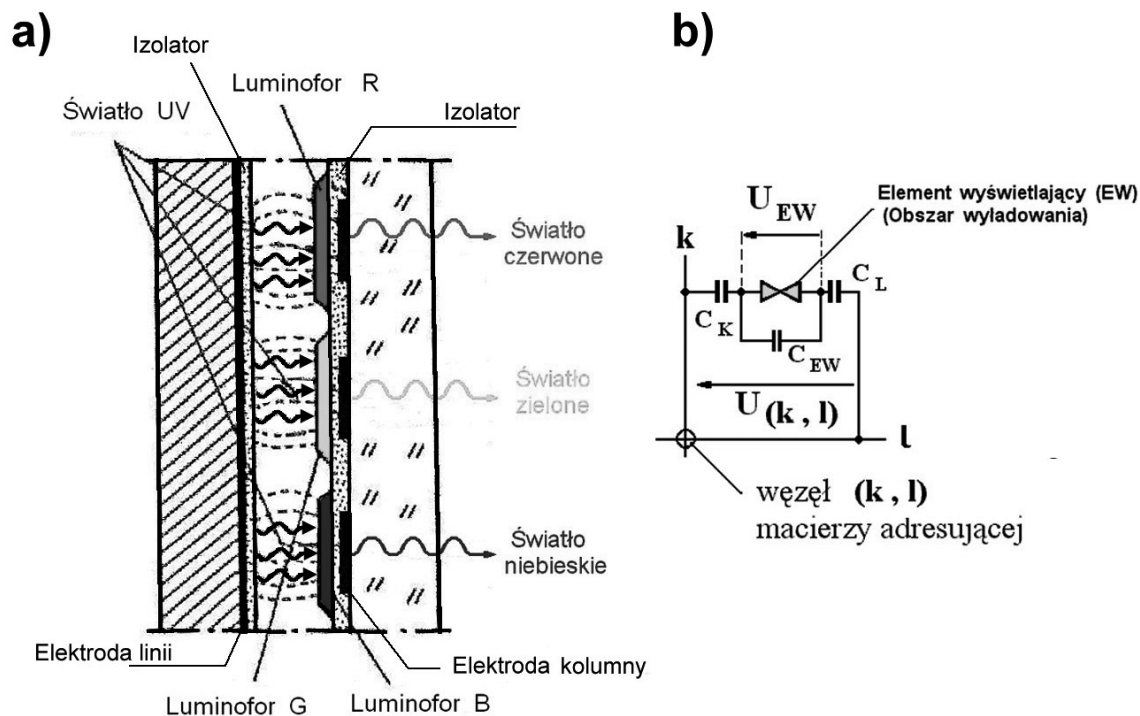


Rysunek 54.

Konstrukcja stałoprądowego wyświetlacza plazmowego



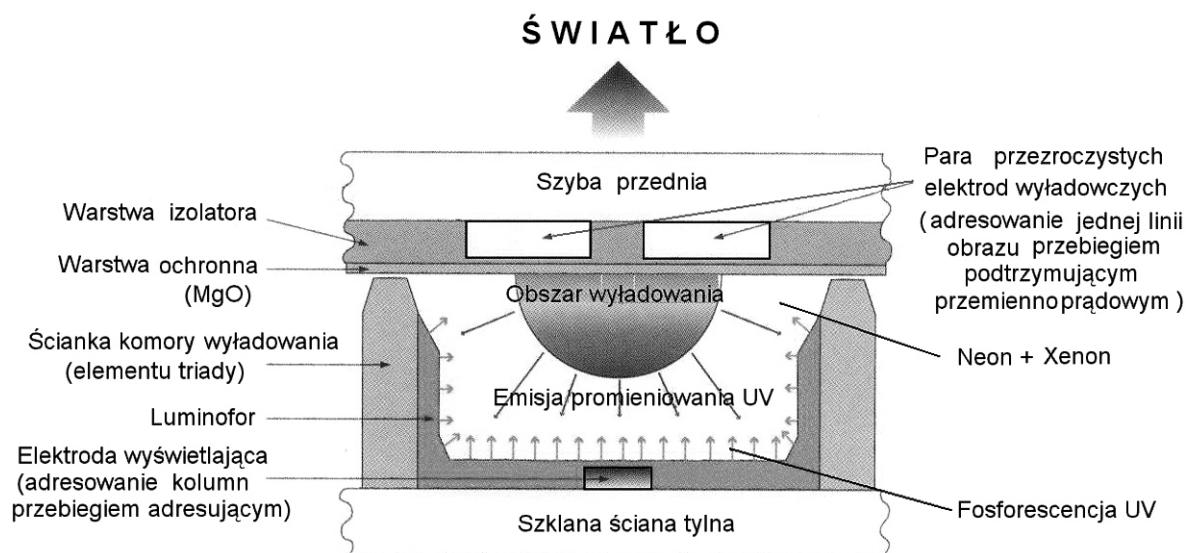
Wad stałoprądowych wyświetlaczy plazmowych nie ma **przebiegnoprądowy wyświetlacz plazmowy**. Istota modyfikacji wobec ekranów DC PDP polega na odizolowaniu elektrod adresujących od gazu. Zapobiega to z jednej strony ich niszczeniu przez jony, z drugiej powoduje wtrącenie do obwodu wyładowania dwóch kondensatorów: C_K i C_L , tworzonych przez elektrody, izolator i obszar wyładowania (rys. 55).



Rysunek 55.

Konstrukcja przebiegnoprądowego wyświetlacza plazmowego

Jeżeli przyjąć, że w chwili początkowej oba te kondensatory nie są naładowane, to po zaadresowaniu węzła napięciem $U(k, l) > U_Z$ (U_Z – napięcie zapłonu) cały potencjał węzła odłoży się na obszarze wyładowania U_{EW} , co spowoduje jego zaświecenie i przepływ w obwodzie wyładowania krótkiego impulsu prądowego, wykładniczo malejącego, ładującego kondensatory C_K i C_L do napięcia $U_Z/2$. Po naładowaniu kondensatorów C_K i C_L prąd w obwodzie wyładowania osiągnie wartość zerową, a obszar wyładowania – po wygenerowaniu krótkiego „błysku” (impulsu świetlnego) – przestanie świecić. Czas trwania „błysku” jest na tyle krótki, że wyładowanie jarzeniowe nie zdąży rozszerzyć się poza obszar adresowanego węzła, co eliminuje konieczność stosowania trudnej do wykonania przegrody izolacyjnej. Ponadto małe pojemności kondensatorów C_K i C_L ograniczają maksymalny ładunek jaki może przepłynąć w obwodzie wyładowania, a w konsekwencji także maksymalną wartość prądu wyładowania, co czyni zbędnym rezystor ograniczający ten prąd. Uzyskiwany impuls świetlny („błysk”) jest zbyt krótki i za słaby, ze względu na małe pojemności kondensatorów C_K i C_L . Problem ten można rozwiązać, zmieniając bezpośrednio po wygenerowaniu „błysku” polaryzację napięcia $U(k, l)$ na przeciwną. Napięcie to doda się do napięć stałych na kondensatorach C_K i C_L , dzięki czemu łączny spadek napięcia na obszarze wyładowania U_{EW} znów przekroczy wartość napięcia zapłonu U_Z i element EW ponownie zacznie świecić do czasu przeładowania kondensatorów C_K i C_L , generując kolejny „błysk”. Zmieniając periodycznie polaryzację $U(k, l)$ z dostatecznie dużą częstotliwością, przez podanie do węzła napięcia przemiennego o częstotliwości rzędu kilkuset kHz i wartości międzyszczytowej równej $2U_Z$, uzyskuje się ciągłą generację impulsów świetlnych z elementu wyświetlającego, którą oko – ze względu na częstotliwość powtarzania błysków, rzędu kilkudziesięciu kHz – odbiera jako ciągłe świecenie. Zmieniając czas dołączenia napięcia przemiennego do węzła można generować „pakiety błysków” o zmiennej długości, sterując w ten sposób jasnością świecenia piksela. Od charakterystycznego sposobu pobudzania obszaru wyładowczego do świecenia wyświetlacze stosujące opisaną wyżej zasadę wyświetlania noszą nazwę przebiegnoprądowych ekranów plazmowych AC PDP. Do tej grupy rozwiązań zaliczają się wszystkie ekrany plazmowe, dostępne obecnie na rynku.



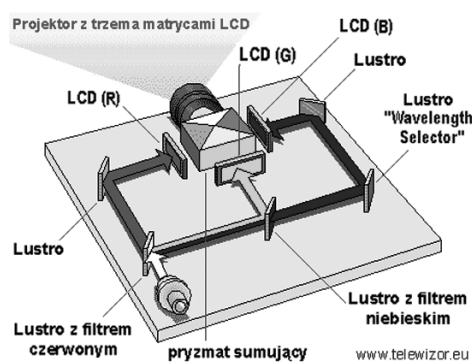
Rysunek 56.

Konstrukcja pojedynczego piksela wyświetlacza plazmowego

Na rys. 56 zilustrowano sposób uzyskiwania obrazów wielobarwnych przez ekrany plazmowe AC PDP. Wnętrze wyświetlacza wypełnione jest gazem lub mieszaniną gazów, świejących podczas wyładowania jarzeniowego światłem nadfioletowym (UV), które pobudza do świecenia paski luminoforów naniesione od wnętrza bańki. Umieszczenie luminoforów na ścieżce wyładowania powodowałoby ich niszczenie (wypalanie) przez jony świejącego gazu. Zjawisko to eliminuje stosowana w obecnie oferowanych rozwiązaniach wielobarwnych wyświetlaczy plazmowych typu AC PDP konstrukcja piksela ich ekranu. Elektrody wyładowcze: kolumn i linii, są tu umiejscowione obok siebie na przedniej szybie piksela, a luminofor rozmieszczony na jego przeciwnej ścianie. Przepływ prądu wyładowania (jonów mieszaniny gazów) odbywa się pomiędzy elektrodami wyładowczymi w dużej odległości od luminoforu, do którego dociera jedynie promieniowanie UV emitowane przez świejącą plazmę.

4.9 PROJEKTORY MULTIMEDIALNE

Projektory 3LCD są budowane najczęściej na podstawie trzech ciekłokrystalicznych matryc oświetlanych przez lampę o dużej mocy. Wytwarzany przez lampę oświetlający biały promień świetlny przepuszczany jest przez filtry odwzorowujące trzy podstawowe barwy – czerwoną, zieloną oraz niebieską. Każdy z tych strumieni jest nakierowany na jeden z paneli ciekłokrystalicznych. Obrazy z poszczególnych paneli są następnie nakładane na siebie i przepuszczane przez obiektyw (rys.57).



Rysunek 57.

Zasada działania projektora 3LCD

Projektory 3LCD – zalety i wady:

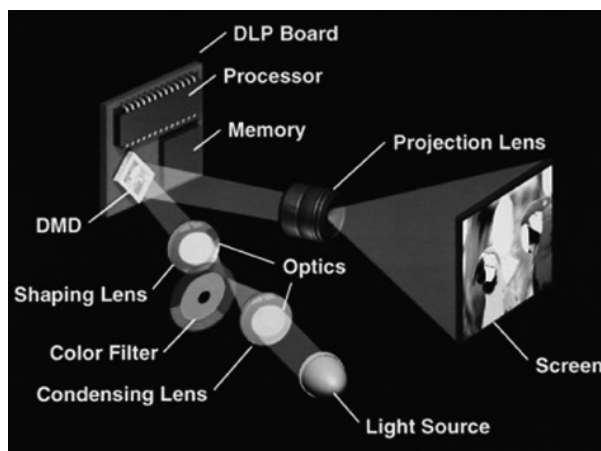
- Kinowe modele projektorów są często wyposażone w panoramiczne matryce o dużych rozdzielczościach przystosowanych do pracy z sygnałem HDTV. Dzięki zwiększeniu rozdzielczości obrazu nie widać pikseli, z których jest tworzony obraz.
- Pojawiły się projektory LCD o bardzo wysokim kontraście. Niemniej ciągłą bolączką rozwiązań opartych na technologii LCD jest niewystarczająca głębia i zróżnicowanie czerni.
- Inne wady: możliwość wystąpienia „martwych pikseli”, w tańszych modelach słaby kontrast, wypalanie paneli w czasie eksploatacji, przy niższych rozdzielczościach widoczne przerwy pomiędzy pikselami
- Zalety: brak efektu tęczy, żywe barwy, bardzo wysoka rozdzielczość najlepszych paneli LCD

Najważniejszą częścią każdego **projektora DLP** jest układ półprzewodnikowy z wbudowanym systemem sterowanych lusterek (DLP chip), wynaleziony przez Dr. Larry Hornbeck z Texas Instruments w 1987 roku. W układ DLP wbudowana jest prostokątna macierz do 2 milionów mikroskopijnych lusterek. Zwierciadła są kwadratami o boku 16 mikronów, a przerwy między nimi nie mogą wynosić więcej niż 1 mikron (rys. 58). Mikrolustra układu DLP umocowane są w specjalnych zawiasach, które umożliwiają im wychylenie się w kierunku źródła światła (ON) albo w odwrotnym (OFF), tworząc w ten sposób jasny albo czarny piksel na powierzchni projekcyjnej. Strumień bitowy zakodowanego obrazu powoduje przelączenie lusterek z pozycji OFF do ON kilka tysięcy razy na sekundę. Kiedy lustro jest w pozycji ON częściej niż OFF odzwierciedla piksel jasnoszary. Lustro, które częściej jest w pozycji OFF odzwierciedla piksel ciemnoszary. W ten sposób lustro w projektorze DLP mogą odzwierciedlić ok.1024 różnych odcieni szarości piksela.



Rysunek 58.
Macierz lusterek

Obraz kolorowy w projektorach DLP tworzony jest w następujący sposób. Białe światło lampy projektora DLP przechodzi przez koło z filtrami kolorów trafia na mikrolustra układu DLP (rys. 59). Koło kolorów rozdziela strumień światła białego na czerwony, zielony i niebieski. Stany włączenia i wyłączenia mikroluster są koordynowane w ten sposób aby tworzyć różne kolory z trzech barw podstawowych. Na przykład lustro, które na utworzyć piksel purpurowy będzie odbijało tylko barwę niebieską i czerwoną. Nasze oczy i mózg łączą te bardzo krótkie, naprzemiennie błyski światła w pożądaną barwę. Układ DLP z kołem kolorów jest w stanie odwzorować w ten sposób przynajmniej 16.7 milionów kolorów. Trójukładowe projektory DLP odwzorowują nie mniej niż 35 trylionów kolorów.



Rysunek 59.
Sposób tworzenia obrazu kolorowego w projektorze DLP z jednym Przetwornikiem



W projektorach najwyższej klasy wyposażonych w 3 przetworniki DLP nie stosuje się już wirującego kolorowego układu optycznego. Każdy przetwornik odpowiada za przetwarzanie jednego z 3 kolorów podstawowych – czerwonego, niebieskiego i zielonego.

Projektory DLP – zalety:

- Największą zaletą zastosowania technologii DLP jest możliwość uzyskania obrazu, który wydaje się **pozbawiony jakichkolwiek łączeń**. Za uzyskanie takiego efektu odpowiada niesamowita bliskość każdego z aluminiowych mikroluster, których 90% powierzchni własnej skutecznie odbija światło w celu wytworzenia obrazu.
- Projektory DLP wyróżniają się **wysokim kontrastem** o współczynniku 2000:1 i większym.
- Projektory DLP mają także **naturalne odwzorowanie barw** i dobrą głębię czerni.
- Zaletą DLP jest także możliwość stosowania **słabszych lamp projekcyjnych** niż w produktach LCD, co ma przełożenie na cichszą pracę układów chłodzących optykę projektora.
- Najnowsze układy DLP umożliwiają tworzenie projektorów **mniejszych gabarytowo, lżejszych i tańszych**.

Projektory DLP – wady:

- Minusem projektorów DLP zbudowanych na podstawie jednego (najczęściej spotykana konfiguracja) mikroprocesora DMD (cyfrowy sterownik mikroluster) jest występowanie tzw. **efektu tęczy** (rys. 60), wynikającego z chwilowego braku zbieżności barw.
- W najprostszyc konstrukcjach filtr składa się z trzech elementów, lecz coraz częściej nawet tanie projektory DLP mają już filtr 4-segmentowy, co zdecydowanie ogranicza wspomniany efekt. Zaawansowane projektory kinowe mają filtr 6-barwny.
- Drugim elementem zmniejszającym efekt tęczy jest zwiększenie prędkości obrotowej wirującego koła. W nowej generacji projektorów DLP pola barwne filtra mają mieć kształt nie prostokątny, lecz "tezkowaty", nieco nachodzący na siebie.



Rysunek 60.
Efekt tęczy

Zadanie 8. Eksport materiału wideo do formatu MPEG-2

W drugiej części ćwiczenia słuchacze zostaną zapoznani z metodami przygotowywania materiału do nagrania płyty DVD oraz authoringu. Program Pinnacle Studio 9 stanowi kompletny zestaw narzędzi umożliwiających tworzenie własnych płyt VCD/SVCD/DVD. Proces ten obejmuje następujące elementy, z którymi słuchacze zostaną zapoznani:

- a) konwersja plików AVI do odpowiednich formatów zgodnych ze standardami zapisu płyt DVD,
- b) budowanie menu scen (authoring),
- c) nagrywanie przygotowanych materiałów na nośnik.

DVD authoring to tworzenie interfejsów płyt DVD, których zawartość można wyświetlać w komputerze osobistym lub odbiorniku telewizyjnym. Opracowany w trakcie edycji interfejs jest narzędziem do sterowania zawartością płyty. Podczas authoring'u wykonywane są następujące czynności:

- zapisanie do postaci cyfrowej obrazu i dźwięku,
- zakodowanie do MPEG-1 lub MPEG-2 obrazu dla DVD-Video,
- zakodowanie do Dolby Digital lub DTS lub MPEG Audio lub LPCM dźwięku dla DVD-Video



- wykonanie projektu graficznego do interfejsu,
- połączenie wszystkich elementów audio-video odpowiednimi linkami w celu zapewnienia współpracy z użytkownikiem,
- sformatowanie obrazu płyty,
- zapisanie tego obrazu na odpowiednim nośniku (np. DVD, dysk twardy) .

Przygotowanie materiału wizyjnego do nagrania na płycie DVD bądź CD składa się z kilku etapów.

Pierwszy etap to zaplanowanie zawartości płyty. Konieczne jest zastanowienie się nad ilością i standardem zapisu ścieżek dźwiękowych, stosowaniem elementów interaktywnych oraz różnymi wersjami językowymi. W etapie planowania konieczne jest również zaplanowanie elementów oraz formy menu. Etap tworzenia menu to wykorzystanie najczęściej gotowego szablonu, który można oczywiście modyfikować. Niektóre programy umożliwiają całkowite przemodelowanie szablonu, zmieniając układ ikon dla rozdziałów, czcionkę w napisach czy tło. Kolejnym etapem jest odpowiedni montaż oraz kompresja wszystkich materiałów, które chcemy wykorzystać do nagrania płyty. Mowa tu przede wszystkim o materiale wizyjnym, ale również o przygotowaniu ścieżek dźwiękowych. Ostatnim etapem jest proces *authoringu*. Termin ten oznacza wszystkie procesy, jakie muszą być zrealizowane po zakodowaniu dźwięku i obrazu, a przed otrzymaniem końcowego produktu. Etap ten polega na połączeniu ze sobą obrazu i wszystkich ścieżek dźwiękowych, podpisów do dialogów, stworzeniu menu oraz podziału na sceny. Korzystając z programu Pinnacle Studio 9 większość tych operacji wykonywane jest automatycznie.

Na etapie eksportu materiału wideo do formatu MPEG-2 słuchacze powinni wykonać następujące operacje:

- zapoznać się z zasadą działania i podstawowymi funkcjami programu Pinnacle Studio 9 dotyczącymi *authoringu*,
- na podstawie założeń przedstawionych podczas ćwiczenia przez prowadzącego przygotować koncepcję tworzonej płyty,
- przeprowadzić konwersję wykorzystywanych materiałów do formatu MPEG-2,
- nagrać gotowy materiał filmowy na dysk komputera lub płytę DVD.

Nagrane filmy jak również inne przygotowane przez prowadzącego należy następnie odtworzyć korzystając z różnych konfiguracji sprzętowych.

Należy zapoznać się z możliwościami, obsługą i konfiguracją monitora, telewizora LCD i rzutnika. Jakim rodzajem rzutnika dysponujemy?

Należy porównać jakość materiału filmowego na monitorze, telewizorze LCD i zaprezentowaną przy użyciu projektora multimedialnego. Co zapewnia najlepszą jakość wyświetlania?

LITERATURA

1. Adobe Premiere Elements. Domowe studio wideo, Helion, Gliwice 2007
2. Beach A., *Kompresja dźwięku i obrazu wideo*, Helion, Gliwice 2009
3. Butryn W., *Dźwięk cyfrowy*, WKiŁ, Warszawa 2002
4. Butryn W., *Dźwięk cyfrowy. Systemy wielokanałowe*, WKiŁ, Warszawa 2004
5. Czyżewski A., *Dźwięk cyfrowy. Wybrane zagadnienia teoretyczne, technologia, zastosowania*, Exit, Warszawa 2001
6. Danowski B., *Komputerowy montaż wideo. Ćwiczenia praktyczne*, Helion, Gliwice 2006
7. Kołodziej P., *Komputerowe studio muzyczne i nie tylko. Przewodnik*, Helion, Gliwice 2007
8. Nasiłowski D., *Jakościowe aspekty kompresji obrazu i dźwięku. Poglądowo o DivX*, Mikom, Warszawa 2004
9. Ozer J., *Tworzenie filmów w Windows XP. Projekty*, Helion, Gliwice 2005
10. Rak R., Skarbata W. (red.), *Wstęp do inżynierii multimedialnych*, Politechnika Warszawska, Warszawa 2004
11. Zieliński T.P., *Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań*, WKiŁ, Warszawa 200



W projekcie **Informatyka +**, poza wykładami i warsztatami, przewidziano następujące działania:

- 24-godzinne kursy dla uczniów w ramach modułów tematycznych
- 24-godzinne kursy metodyczne dla nauczycieli, przygotowujące do pracy z uczniem zdolnym
- nagrania 60 wykładów informatycznych, prowadzonych przez wybitnych specjalistów i nauczycieli akademickich
 - konkursy dla uczniów, trzy w ciągu roku
 - udział uczniów w pracach kół naukowych
 - udział uczniów w konferencjach naukowych
 - obozy wypoczynkowo-naukowe.

Szczegółowe informacje znajdują się na stronie projektu

www.informatykaplus.edu.pl

