

informatyka+

Algorytmika i programowanie

Bazy danych

Multimedia, grafika i technologie internetowe

Sieci komputerowe

Tendencje w rozwoju informatyki i jej zastosowań

informatyka+

Kuźnia Talentów Informatycznych: Multimedia, grafika i technologie internetowe Nagrywanie i obróbka filmów

Andrzej Majkowski

Człowiek – najlepsza inwestycja

Człowiek – najlepsza inwestycja



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.



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.

Nagrywanie i obróbka filmów



Rodzaj zajęć: Kuźnia Talentów Informatycznych

Tytuł: Nagrywanie i obróbka filmów

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 2011

Copyright © Warszawska Wyższa Szkoła Informatyki 2009

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.

Nagrywanie i obróbka filmów



Andrzej Majkowski

Politechnika Warszawska

amajk@ee.pw.edu.pl



Streszczenie

Zakres materiału tego kursu obejmuje następujące zagadnienia:

1. Rejestracja filmów z użyciem różnych narzędzi (różne typy kamer, aparat fotograficzny z funkcją nagrywania, telefon komórkowy). Formaty zapisu filmów. Rejestracja obrazów statycznych i dźwięków jako elementów wzbogacających film. Sposoby przenoszenia filmów z kamery do komputera. Poznanie oprogramowania do przechwytywania plików filmowych z kamery i do wyświetlania filmów.
2. Techniki kodowania obrazów ruchomych DV, MPEG, XVID, DIVX – zalety i wady tych typów kodowania. Konwersja między różnymi formami kodowania filmów i dlaczego musimy ją stosować. Jakość zapisu filmu zakodowanego w różnych formatach. Sposoby dogrywania i łączenia z obrazem dźwięku i napisów.
3. Nieliniowy montaż wideo. Poznanie oprogramowania służącego do montażu nieliniowego. Poprawa jakości filmów, wzbogacanie za pomocą technik grafiki komputerowej, zastosowanie filtrów. Dodawanie menu, opisów, znaczników – authoring.
4. Nagrywanie filmów na płyty DVD. Porównanie jakości z nośnikiem Blu-ray. Ocena jakości filmów prezentowanych na różnego rodzaju wyświetlaczach LCD, plazmowy, projektor (rola ekranu).

Sprzęt i oprogramowanie potrzebne do przeprowadzenia kursu:

1. komputer klasy PC, system Windows XP, nagrywarka DVD, odtwarzacz Blu-ray,
2. kamera pracująca w standardzie mini DV, full HD, kamera internetowa,
3. mikrofon,
4. cyfrowy aparat fotograficzny z funkcją nagrywania filmów, telefon komórkowy z funkcją nagrywania filmów,
5. oprogramowanie np. Pinnacle Studio, Adobe Premiere – w wersji bardziej zaawansowanej,
6. telewizor CRT, LCD, rzutnik multimedialny, odtwarzacz DVD, Blu-ray,
7. zestaw kina domowego.



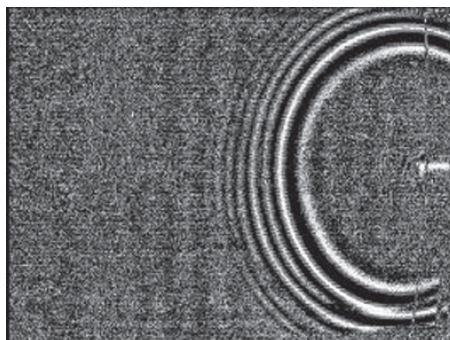
Spis treści

1. Rejestracja dźwięku i obrazu	6
1.1. Dźwięk	6
1.1.1. Jak odbieramy dźwięk	6
1.1.2. Zakres słyszalności	7
1.1.3. Ocena jakości dźwięku	8
1.2. Obraz	8
1.2.1. Modele barw.....	10
1.2.2. Kontrast, korelacja gamma, temperatura barwowa, balans bieli.....	11
1.3. Rejestracja sygnałów dźwiękowych.....	14
1.3.1. Podział akustyczny mikrofonów.....	14
1.3.2. Podział mikrofonów ze względu na rodzaj przetwornika	15
1.3.3. Parametry mikrofonów	16
1.3.4. Te Chinki mikrofonowe	17
1.4. Formaty zapisu i przechowywania plików multimedialnych.....	17
1.5. Analizatory obrazów.....	19
Zadanie 1. Pozyskiwanie materiału wideo	24
2. Przetwarzanie sekwencyjne wideo	26
2.1. Komputerowa edycja obrazu.....	26
2.2. Standardy kodowania sekwencji wideo MPEG	27
2.2.1. Przebieg procesu kodowania w standardach MPEG-1 – MPEG-2	31
2.2.2. Różnice pomiędzy standardami MPEG-1 i MPEG-2	33
2.2.3. Kompresja MPEG-4	35
Zadanie 2. Zgrywanie sekwencji wizyjnej	37
3. Edycja i montaż obrazów wideo	38
3.1. Telewizja analogowa i cyfrowa.....	38
3.1.1. Standard telewizji kolorowej HDTV.....	39
3.1.2. Cyfrowa telewizja systemu DVB	39
3.2. Poprawa jakości obrazu.....	39
3.2.1. Redukcja artefaktów wynikających z kompresji	40
3.2.2. Eksponowanie konturów obrazów	40
3.2.3. Algorytmy poprawy jakości obrazu.....	41
3.3. Komputerowa edycja obrazu	43
3.3.1. Authoring VCD i DVD	44
Zadanie 3. Edycja sekwencji wizyjnej oraz eksport materiału wideo do formatu MPEG-2.....	46
4. Odtwarzanie treści multimedialnych	48
4.1. Kino domowe.....	48
4.2. Akustyka pomieszczenia	50
4.3. Wyświetlacze LCD.....	52
4.4. Ekrany plazmowe	57
4.5. Projekторы multimedialne	60
Zadanie 4: Eksport materiału wideo do formatu MPEG-2	62
Literatura	63

1 REJESTRACJA DŹWIĘKU I OBRAZU

1.1 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, iż cząsteczka pobudzona przekazuje energię cząstce 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.



Rysunek 1.
Fala dźwiękowa

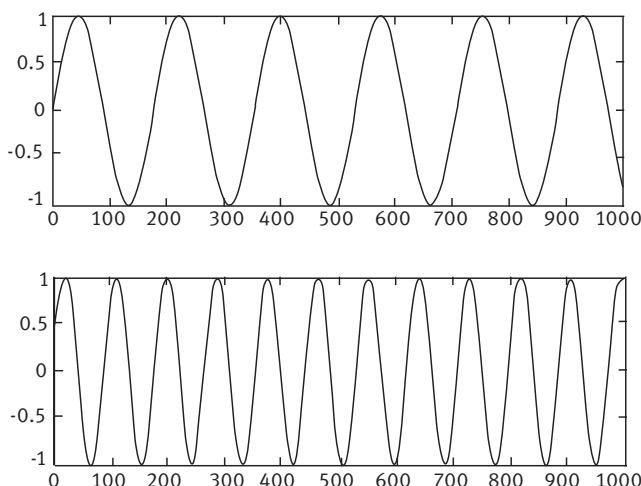
Zakres częstotliwości od 20 Hz do 20 kHz jest zakresem częstotliwości słyszalnych (fonicznych, audio). Dźwięki o częstotliwości mniejszej od 20 Hz są nazywane **infradźwiękami**, zaś o częstotliwości większej od 20k Hz – **ultradźwiękami**.

1.1.1 JAK ODBIERAMY DŹWIĘKI

Elementarnym rodzajem dźwięku, dla którego fala dźwiękowa ma postać sinusoidy (rys. 2), jest **ton**. **Wysokość tonu** to atrybut wrażenia słuchowego, umożliwiający uszeregowanie dźwięków na skali niskie-wysokie. Przez **wysokość dźwięku** rozumie się częstotliwość drgań fali akustycznej, im wyższa częstotliwość drgań tym „wyższy” dźwięk. Na rys. 2 częstotliwość drugiego sygnału jest dwa razy większa niż pierwszego, zatem dźwięk o takim przebiegu będzie odbierany jako „wyższy”. Dźwięki są najczęściej sygnałami złożonymi (występuje w nich wiele składowych sinusoidalnych o różnych amplitudach i częstotliwościach). Wysokość dźwięku, 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.

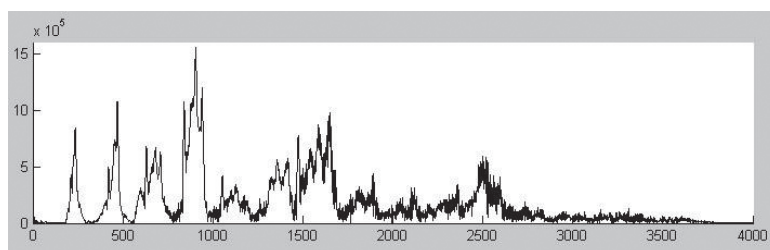
Z pojęciem wysokości dźwięku związane są **interwały muzyczne**, czyli „odległości” między dźwiękami na skali muzycznej. Określone są stosunkiem częstotliwości sygnałów. Oktawa jest to interwał określający dźwięki, których stosunek częstotliwości jest równy 2:1. Człowiek jest w stanie interpretować poprawnie interwały muzyczne dla tonów o częstotliwości max ok. 5 kHz. Powyżej 2,5 kHz występują znaczne błędy. Natomiast powyżej częstotliwości 5 kHz występuje brak wrażenia melodii, chociaż spostrzegane są różnice częstotliwości.

Bardzo często w analizie sygnału dźwiękowego korzysta się z jego częstotliwościowej reprezentacji. Mówimy wtedy o tzw. **widmie** sygnału dźwiękowego. Widmo sygnału dźwiękowego umożliwia zobrazowanie składowych sinusoidalnych, będących funkcjami czasu (ich częstotliwości i amplitudy), tworzących dany dźwięk. Na rys. 3 przedstawione jest przykładowe widmo sygnału dźwiękowego. Na osi x są umieszczone częstotliwości składowych sinusoidalnych. W tym przypadku zawierają się one w zakresie od 43 Hz do 12000 Hz. Z osi y można odczytać pośrednio informację o amplitudach składowych sinusoidalnych.



Rysunek 2.

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



Rysunek 3.

Widmo sygnału dźwiękowego

Barwa dźwięku to cecha wrażenia słuchowego, dzięki której rozróżniamy dźwięki o tej samej głośności i częstotliwości. Barwa dźwięku zależy głównie od jego struktury widmowej, natężenia dźwięku i przebiegu czasowego dźwięku. I tak interesujące eksperymenty pokazują, że w przypadku niektórych instrumentów ważniejszą rolę odgrywa struktura widmowa (klarnet, trąbka), a innych – czasowa (flet). Kluczową rolę odgrywa też proces narastania i trwania dźwięku.

Słuch ludzki charakteryzuje pewna niesymetryczność w odbiorze wysokości dźwięków w uchu lewym i prawym. U zdrowego człowieka różnice nie przekraczają zwykle 3%. Osoby o słuchu muzycznym potrafią określić wysokość dźwięku z dokładnością do 0,3-1%.

1.1.2 ZAKRES SŁYSZALNOŚCI

Głośność to taka cecha wrażenia słuchowego, która umożliwia uszeregowanie dźwięków na skali głośno-cicho. Teoretycznie ucho ludzkie potrafi odebrać i przetworzyć drgania o częstotliwości od 16 Hz do 20 kHz. Jest to jednak duże uproszczenie, nie mające wiele wspólnego z rzeczywistością. Okazuje się, że ten zakres jest słyszalny tylko wtedy, gdy energia dźwięku jest duża. Przy cichych dźwiękach czułość ucha drastycznie maleje w obszarze częstotliwości poniżej 200 Hz oraz powyżej 8 kHz. 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 sek. do ok. 3 min. 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 1000 Hz) wykrywanego przez ucho ludzkie wynosi średnio 20 μPa (rys. 4). **Próg bólu** jest to wartość ciśnienia akustycznego, przy której ucho odczuwa wrażenie bólu. Jest ono prawie niezależne od częstotliwości i wynosi 140 dB dla dźwięków sinusoidalnych oraz 120 dB 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.



Rysunek 4.
Zakres słyszalności człowieka

Okazuje się, że człowiek nie wszystkie dźwięki o tym samym poziomie głośności słyszy jednakowo dobrze. Dźwięki bardzo niskie i bardzo wysokie są słyszane słabo, za to tony o częstotliwościach od 1 KHz do 5 KHz (mniej więcej zakres mowy ludzkiej) są słyszane wyjątkowo dobrze. Na przykład ton 10 dB o częstotliwości 1000 Hz będzie przez większość osób świetnie słyszalny, ale ton 10 dB o częstotliwości 25 Hz chyba wszyscy usłyszymy 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.

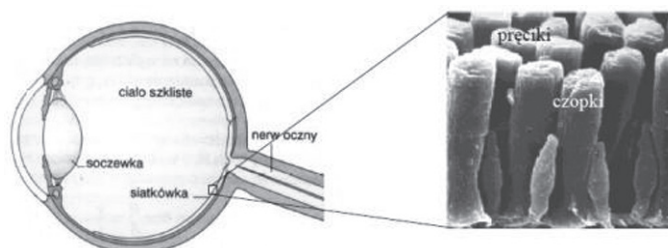
1.1.3 OCENA JAKOŚCI DŹWIĘKU

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

1.2 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 niezdolność 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

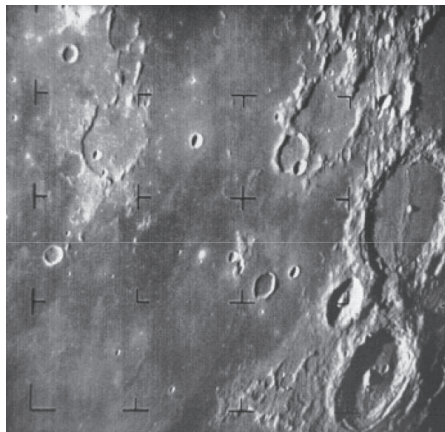
Obrazy cyfrowe reprezentują te same sceny, które możemy obserwować, ale przedstawione w postaci dwuwymiarowych tablic pikseli. Technika cyfrowa umożliwia przeprowadzenie wielu operacji obróbki obrazu, w tym także działań niewykonalnych tradycyjnymi metodami przy pomocy szklanych filtrów optycznych lub analogowej elektroniki. Jedną z pierwszych prób wykorzystania techniki cyfrowej w praktyce było przesyłanie obrazów na odległość z wykorzystaniem kabla. Do przesyłania 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 przesyłanie 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.

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).



Rysunek 6.
Obraz przetransmitowany i odtworzony przez Bartlane System

Na rys. 7 przedstawiono obraz Księżyca sfotografowany przez statek Ranger 7. Zdjęcie zostało wykonane w 1964 roku przy użyciu kamery telewizyjnej i następnie przesłane na Ziemię. Zdjęcia 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

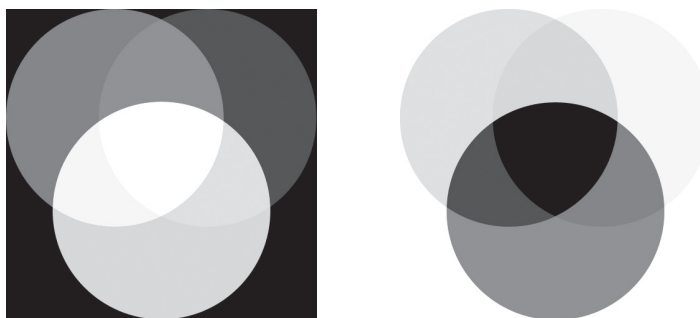
1.2.1 MODELE BARW

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 to model barw ukierunkowany 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 to model ukierunkowany 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).



a)

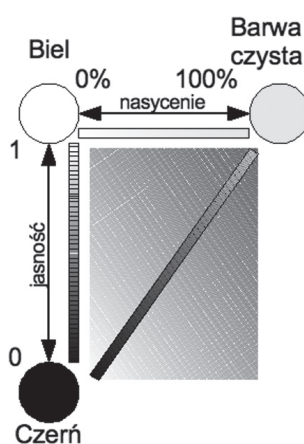
b)

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 – świecących i odpowiada wrażeniu słabszego lub mocniejszego strumienia światła.

Odcień barwy, jasność i nasycenie (trzy atrybuty barwy) są ze sobą ściśle związane. Zmiana jednego atrybutu pociąga za sobą zmianę pozostałych (rys. 9). W zakresie widzenia barwnego wraz ze zmianą jasności zachodzą zmiany barwy postrzeganej. Wrażenie zmiany barwy obserwujemy również, gdy bez zmiany odcienia i jasności zmniejszymy nasycenie barwy.



Rysunek 9.

Atrybuty barwy

1.2.2 KONTRAST, KOREKCJA GAMMA, TEMPERATURA BARWOWA, BALANS BIELI

Przy opisie przetwarzania obrazów rejestrowanych aparatami cyfrowymi, czy kamerami cyfrowymi często używa się pewnych podstawowych pojęć. Poznanie znaczenia tych pojęć umożliwia lepsze zrozumienie bar-

dziej złożonych procesów zachodzących podczas przetwarzania obrazów cyfrowych. Pojęcia te to: kontrast, korekcja gamma, temperatura barwowa, balans bieli.

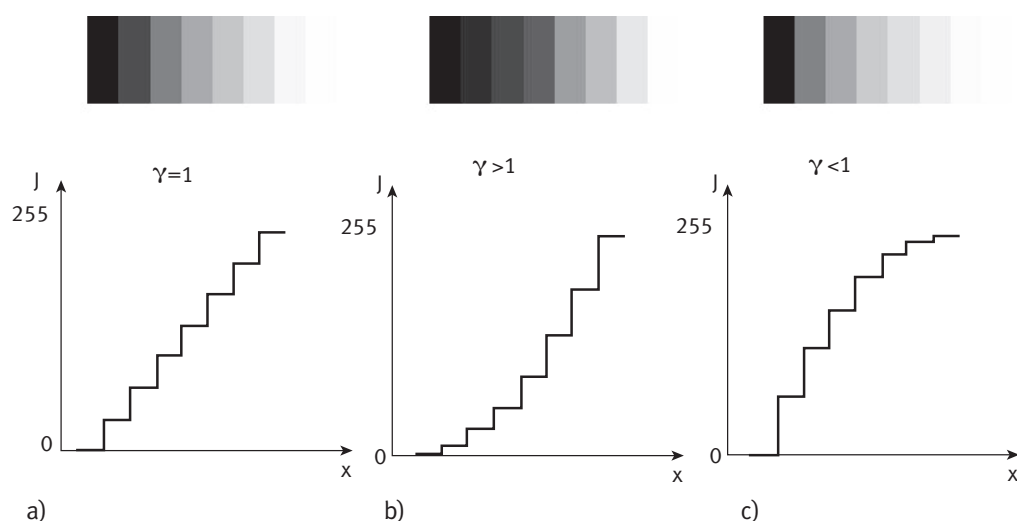


Rysunek 10.

Efekt zastosowania korekcji gamma, lewy górny róg – obraz oryginalny, pozostałe obrazy są wynikiem zastosowania korekcji gamma z różnym współczynnikiem γ

Kontrast określa zróżnicowanie jasności poszczególnych punktów ekranu. Z punktu widzenia optymalnej reprodukcji obrazu nie jest tylko istotny maksymalny stosunek pomiędzy najjaśniejszym i najciemniejszym fragmentem ekranu (kontrast maksymalny), lecz także rozkład różnic w jasności poszczególnych części obrazu (gradacja kontrastu). Dla osiągnięcia wiernej reprodukcji rzeczywistości charakterystyka jasności układu przetwarzania i wyświetlania obrazu powinna być liniowa (rys. 11a). Z subiektywnego punktu widzenia niekiedy wskazane jest specjalne kształtowanie gradacji kontrastu. Często stosuje się nieliniowe przetwarzanie, np. w celu pełnego wykorzystania dynamiki obrazu (czyli poprawnego zróżnicowanie skali szarości zarówno w jasnych jak i ciemnych partiach obrazu – rys. 10). Technika cyfrowa daje tutaj możliwości nieosiągalne dla techniki analogowej.

Często wprowadza się celowo pewną nieliniowość przetwarzania, aby w efekcie otrzymać liniową charakterystykę końcową. W przypadku liniowej charakterystyki przetwarzania (rys. 11a) jasność obrazu J jest proporcjonalna do czynnika ją wywołującego (np. napięcia x na przetworniku). Nieliniowa charakterystyka świetlna $J-x$ (rys. 11b, 11c) może być opisana w następujący sposób $J \sim x^\gamma$, czyli jasność obrazu jest proporcjonalna do wywołującego ją napięcia x podniesionego do potęgi γ . Wykładnik γ oznacza stopień nieliniowości przetwornika. Od greckiej litery γ określającej ten współczynnik, korekcja charakterystyki przeprowadzana w ten sposób nosi nazwę **korekcji gamma**.

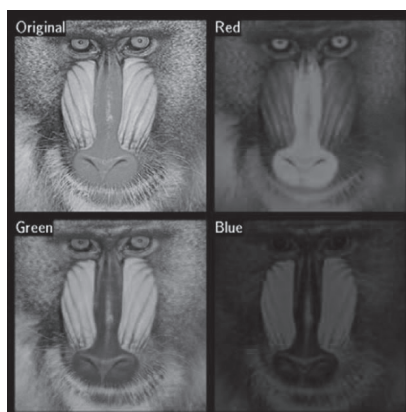


Rysunek 11.
Ilustracja korekcji gamma

W systemie przetwarzania i wyświetlania obrazów istotną rzeczą jest wierna reprodukcja barw. Barwa obiektów zarejestrowana przez kamerę czy aparat fotograficzny zależy od koloru oświetlenia. W tym przypadku, barwa, np. koloru skóry czy bieli śniegu na zdjęciu, może być różna od tej jakiej oczekujemy. Zadaniem korekcji barw jest właśnie sprowadzenie postaci barw do postaci akceptowalnej przez widza. Prawidłowe odwzorowanie koloru śniegu jest przykładem ustawienia **balansu bieli**. Często na odbitkach fotograficznych wykonanych z tego samego negatywu w różnych zakładach fotograficznych są widoczne różnice w jego zabarwieniu: śnieg przybiera zabarwienie niebieskie, żółte, zielone, a niekiedy różowe. Celem ustawienia balansu bieli jest usunięcie tego zabarwienia.

Niekiedy potrzebne jest połączenie procesu korekcji barw z korekcją gamma dla obrazu czy sygnału wizyjnego rozłożonego na składowe RGB (rys. 12). Korekcje stosuje się oddzielnie dla każdego obrazu: czerwonego, zielonego i niebieskiego. Inaczej mogą się pojawić zafaszowania barw w zależności od jasności poszczególnych fragmentów obrazu. Jest to wynik tzw. braku równowagi dynamicznej bieli.

Temperatura barwowa, jako cecha określająca wrażenie percepcyjne oglądanego obrazu, zależy głównie od rodzaju oświetlenia oraz od właściwości barwnych elementów występujących w scenie obrazowej. W praktyce temperaturę barwową definiuje się na podstawie relacji, jakie zaobserwowano pomiędzy temperaturą a właściwościami emisyjnymi ciała czarnego.



Rysunek 12.
Obraz rozłożony na składowe RGB (czerwony, zielony, niebieski)



Temperaturę barwową oblicza się na podstawie średniej wartości kolorów całego obrazu, z pominięciem pikseli, które nie mają wielkiego wpływu na temperaturę barwową, a mianowicie pikseli koloru czarnego i tzw. pikseli samo-świejących, czyli o jasności większej od wartości średniej o pewną wartość progową. Obraz kwalifikowany jest do kategorii barwowej według przedziału temperatur, do którego należy obliczona wartość. Przedziały te zostały wyznaczone doświadczalnie za pomocą badań subiektywnych (patrz tabela poniżej).

Kategoria subiektywna	Zakres temperatur
Gorąca	1667K ~ 2250K
Ciepła	2251K ~ 4170K
Neutralna	4171K ~ 8060K
Zimna	8061K ~ 25000K



Rysunek 13.

Przykład zdjęć o różnej temperaturze barwowej

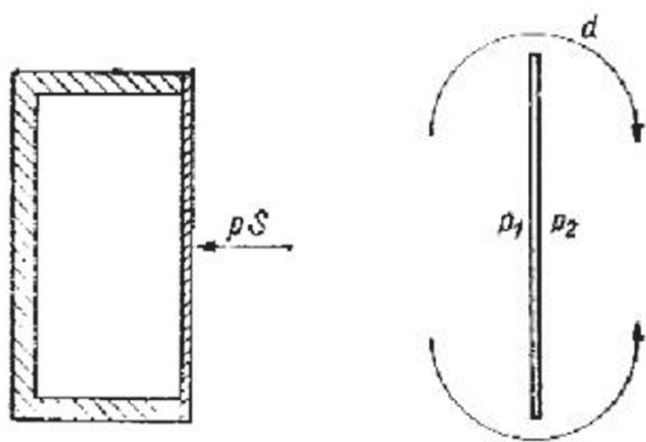
Na rys. 13 przedstawiono trzy zdjęcia, których temperatura barwowa jest różna. Zdjęcie z lewej na neutralną temperaturę barwową, w środku – temperatura barwowa jest przesunięta ku czerwieni, zdjęcie z prawej ma temperaturę barwową przesuniętą w stronę barwy niebieskiej.

1.3 REJESTRACJA SYGNAŁÓW DŹWIĘKOWYCH

Aby możliwa była reprodukcja dźwięku, w większości przypadków jest konieczna najpierw rejestracja dźwięku. Do tego celu stosuje się mikrofony różnego typu, 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.

1.3.1 PODZIAŁ AKUSTYCZNY MIKROFONÓW

- **Mikrofony ciśnieniowe** są wszechkierunkowe. Membrana w tych mikrofonach jest otwarta tylko z jednej strony (rys. 14a). 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. 14b).
- **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.



a)

b)

Rysunek 14.

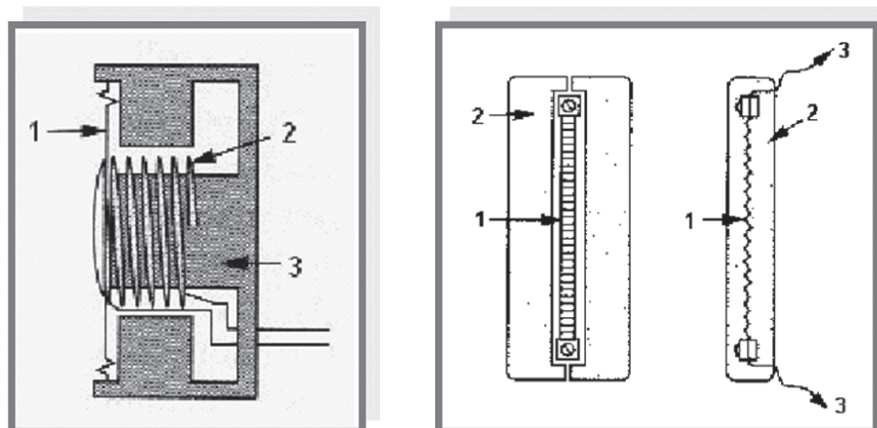
Zasada działania mikrofonu ciśnieniowego a) i gradientowego b)

1.3.2 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. 15a).
- **wstęgowe** – padająca fala akustyczna pobudza do drgań wstęgę umieszczoną w polu magnetycznym, co powoduje powstanie proporcjonalnego do niej sygnału elektrycznego (rys. 15b).

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).



a)

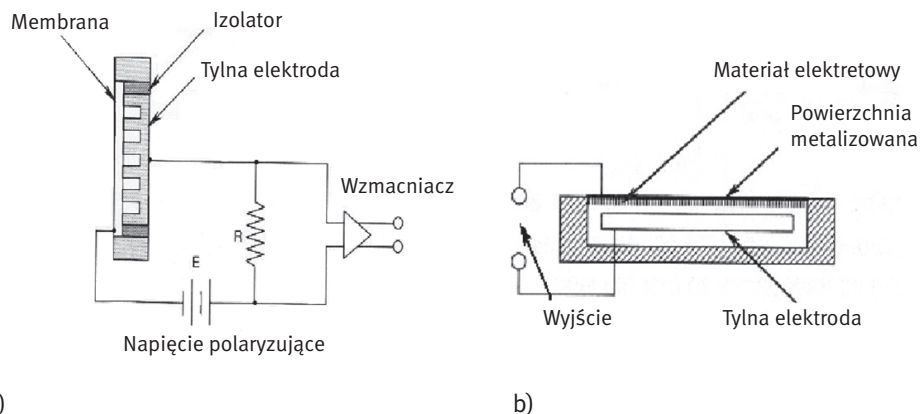
b)

Rysunek 15.

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

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

Mikrofony pojemnościowe (rys. 16a) mają małe wymiary i masę, prostą budowę i małą wrażliwość na wstrząsy. Wychylenia membrany tych mikrofonów są bardzo małe – dochodzą najwyżej do kilku tysięcznych 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. 16b). Mikrofony pojemnościowe są powszechnie stosowane w przenośnych urządzeniach do nagrywania dźwięków, np. w magnetofonach, videokamerach itp.



Rysunek 16.

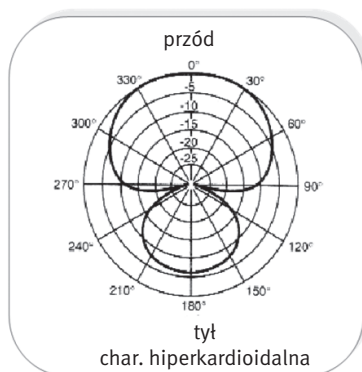
Zasada działania mikrofonu pojemnościowego a) i elektretowego b)

1.3.3 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. 17). 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.

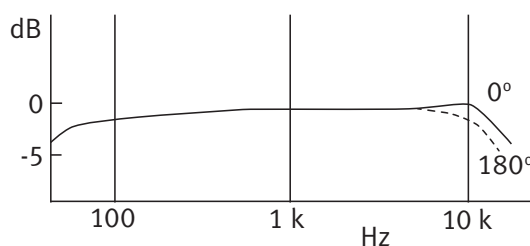
Z kolei **charakterystyka widmowa** określa poziom sygnału na wyjściu mikrofonu w funkcji częstotliwości (rys. 18). Najbardziej wierne przetwarzanie sygnału uzyskuje się w przypadku płaskiej charakterystyki widmowej. Często jednak mikrofony są dedykowane do przetwarzania konkretnych dźwięków, np. mowy, instrumentów perkusyjnych itd. Mikrofony mają wtedy specjalnie dobrane charakterystyki widmowe, tak aby najlepiej wydobyć w sygnale pożądane częstotliwości.



Rysunek 17.

Przykładowa hiperkardioidalna charakterystyka kierunkowa mikrofonu

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.



Rysunek 18.

Przykładowa charakterystyka widmowa mikrofonu

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

1.3.4 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 kardiodalnej 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.

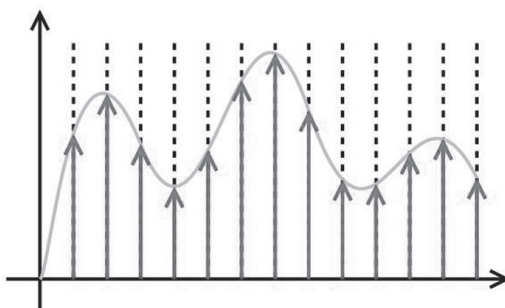
Technika MS (ang. *Middle and Side*) – w tym przypadku używa się dwóch mikrofonów: jednego o charakterystyce kardiodalnej, drugiego o ósemkowej lub dwóch mikrofonów o charakterystyce ósemkowej. Mikrofon *M* (*middle*, kardioda lub ósemka) zbiera dźwięk bezpośredni, zaś *S* (*side*, ósemka) – dźwięk odbity 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.

1.4 FORMATY ZAPISU I PRZECHOWYWANIA PLIKÓW MULTIMEDIALNYCH

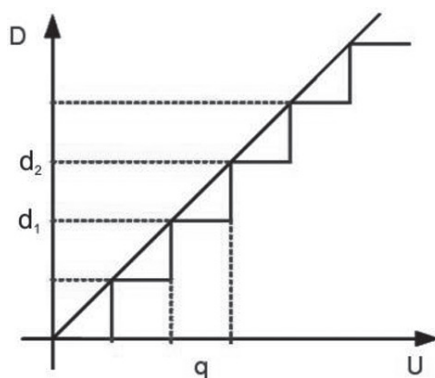
Pliki przechowujące materiały multimedialne często muszą umożliwić zapis i przechowywanie różnego rodzaju danych: dźwięku, obrazów, filmu, napisów, informacji o rozdziałach itp. Potrzebny jest do tego specjalny format zapisu danych, który będzie umożliwiał poprawne wyświetlenie lub synchronizację danych w celu ich jednoczesnego odtworzenia. Ogólnie taki format zapisu nazywa się **kontenerem multimedialnym**. Dźwięk przechowywany w kontenerze multimedialnym musi być zapisany w jakiejś postaci cyfrowej. Jedną z najpo-

pularniejszych 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 (próbki) w określonych (najczęściej równych) odstępach czasu (rys. 19), czyli z określoną częstotliwością (tzw. częstotliwość próbkowania).



Rysunek 19.
Ilustracja zapisu dźwięku w formacie PCM

Wartość chwilowa sygnału jest przedstawiana za pomocą słowa kodowego, którego wartości odpowiadają 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. 20. 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 20.
Kwantyzacja sygnału

Dźwięk w formacie PCM może być zapisywany z różną częstotliwością próbkowania, najczęściej jest to 8 kHz (niektóre standardy telefonii), 44.1 kHz (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 dokólna). Reprezentacja dźwięku próbkowana z częstotliwością 44.1 kHz i w rozdzielczości 16 bitów na próbkę (65536 możliwych wartości amplitudy fali dźwiękowej na próbkę) jest uwa-

żana 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 oryginalną. 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 pozostaje 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 1997 roku. Format AAC został zaprojektowany jako następcą MP3, oferujący lepszą jakość dźwięku przy podobnym rozmiarze danych. Kompresja AAC jest modularna i oferuje w standardowo cztery profile:

- Low Complexity (LC) – najprostszy, najszerszej 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 dźwięku

- próbkowanie 8-96 kHz (MP3 16-48 kHz);
- do 48 kanałów (MP3 – 2 kanały w standardzie MPEG-1 i 5.1 w standardzie MPEG-2);
- skuteczniejszy i wydajniejszy;
- lepsze przenoszenie częstotliwości ponad 16 kHz,
- lepszy tryb kompresji sygnału stereofonicznego joint-stereo.

1.5 ANALIZATORY OBRAZÓW

Przetworniki obrazów świetlnych można podzielić na:

- **analizatory obrazów**, które wykonują złożony proces zamiany przebiegu wejściowego (czyli np. tego, co widzi obiektyw kamery) na elektryczny sygnał wyjściowy, będący prostą funkcją tylko jednego parametru – czasu,



- **wyświetlacze obrazów** (np. monitory, telewizory), które wytwarzają niejako „produkt końcowy” całego procesu przetwarzania obrazów, podlegający fizycznej weryfikacji przez obserwatora, który tylko na tej podstawie może ocenić jakość tego procesu.

W dalszej części zajmiemy się analizatorami obrazów, które stanowią pierwszy element toru cyfrowego pozyskiwania obrazu.

Analizator obrazów świetlnych realizuje konwersję obrazów świetlnych na odpowiadający im sygnał elektryczny. Przetwornikiem dokonującym takiej konwersji może być fotokomórka, bądź fototranzystor. Analizatory obrazów wyróżnia jednoczesne zachodzenie w nich, dla każdego elementu analizowanego obrazu, trzech procesów: **przetwarzania optoelektrycznego**, polegającego na proporcjonalnej do natężenia oświetlenia modyfikacji elektrycznych właściwości ciała stałego, **akumulacji** wytworzonej informacji elektrycznej w miejscu jej powstania, do czasu jej odczytu oraz **adresowania**, czyli odczytania wytworzonej informacji elektrycznej i opatrzenia jej adresem, umożliwiającym określenie miejsca padania odpowiadającego tej informacji strumienia świetlnego.

Przetwarzanie optoelektryczne jest oparte na **efekcie fotoelektrycznym**. Efekt fotoelektryczny polega na uwalnianiu elektronów pozostających normalnie w stanie niewzbudzonym (w tzw. paśmie podstawowym) do paśmie przewodzenia, w wyniku absorpcji przez elektron fotonu. Wyróżniamy **efekt fotoelektryczny zewnętrzny**, jeżeli energia fotonu jest na tyle duża, że elektron po wzbudzeniu opuści strukturę materiału oraz **efekt fotoelektryczny wewnętrzny**, jeżeli wzbudzony elektron pozostaje wewnątrz struktury materiału. W analizatorach obrazu jest wykorzystywany efekt fotoelektryczny wewnętrzny.

Materiał wykazujący zjawisko efektu fotoelektrycznego wewnętrznego, zwany dalej **materiałem światłoczułym** (fotoprzewodnikiem), jest zawsze półprzewodnikiem o dużej rezystancji. Jednostronne oświetlenie materiału światłoczułego strumieniem świetlnym, pochodzącym od analizowanego obrazu spowoduje pojawienie się wolnych ładunków w paśmie przewodzenia materiału, w ilości proporcjonalnej do natężenia oświetlenia danego miejsca płytki. Oznacza to, że po naświetleniu płytki w jej strukturze powstaje ładunkowa replika padającego obrazu świetlnego.

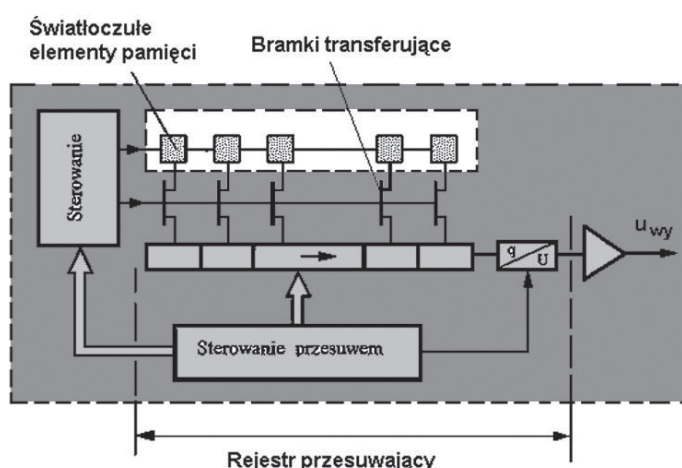
Pojawienie się w strukturze płytki dodatkowych nośników prądu zmienia także jej fizyczne właściwości, w szczególności powoduje spadek rezystancji skrośnej, proporcjonalny do liczby uwolnionych nośników. Po naświetleniu płytki w jej strukturze powstaje zatem druga – rezystancyjna replika padającego obrazu świetlnego. Obie wymienione elektryczne repliki analizowanego obrazu świetlnego są wykorzystywane we współczesnych analizatorach obrazu powszechnego zastosowania.

Akumulacja ładunku ma za zadanie podwyższenie czułości analizatora obrazów. Podczas analizy obrazu, tj. sekwencyjnego odczytywania wartości fotoładunku odpowiadającemu każdemu z pikseli, czas dostępu do każdego z nich jest bardzo krótki. Wygenerowany w tym czasie fotoładunek jest za mały, aby wytworzony za jego pomocą wyjściowy prąd przetwornika miał satysfakcjonującą wartość przy normalnie stosowanych warunkach oświetlenia przetwornika. Akumulacja fotoładunków umożliwia ich gromadzenie w strukturze materiału światłoczułego piksela także pomiędzy jego odczytami, dzięki czemu liczba nośników prądu w momencie odczytu jest wielokrotnie większa. W rezultacie proporcjonalnie rośnie wyjściowy prąd sygnału analizatora, a więc w efekcie – jego czułość.

Jeżeli właściwości przewodzące materiału światłoczułego, z którego jest wykonana płytka są słabe (bliskie izolatorowi), akumulację zapewnia sam materiał światłoczuły, ponieważ duża wartość jego rezystancji powierzchniowej uniemożliwia rozptył wytworzonych w danym pikselu fotoładunków do sąsiednich pikseli. W przypadku materiałów o relatywnie niskiej rezystancji powierzchniowej, jak ma to miejsce w przypadku najpopularniejszego obecnie materiału światłoczułego, jakim obecnie jest monokrystaliczny krzem, akumulację zapewniają kondensatory.

Następnie odczytuje się zapisaną informację elektryczną. W celu określenia miejsca padania odpowiadającego tej informacji strumienia świetlnego stosuje się różne metody adresowania (określenia położenie poszczególnych pikseli). Często stosowany sposób adresowania, przedstawiony na rys. 21, wykorzystuje re-





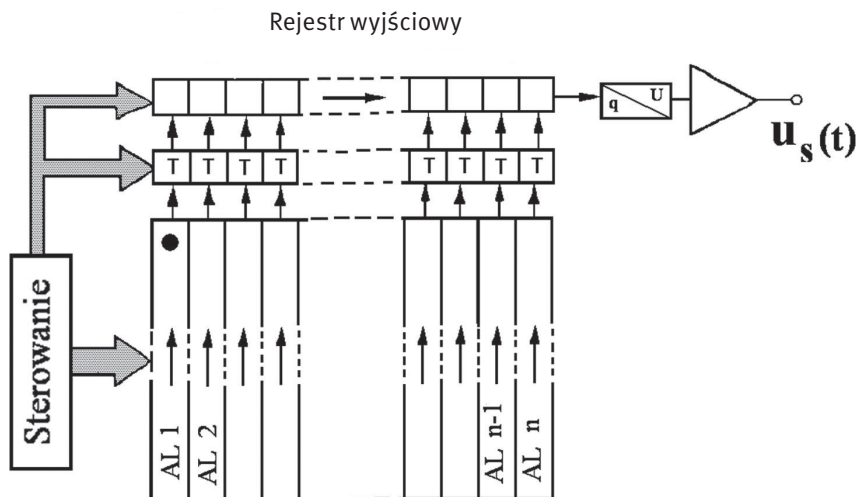
Rysunek 21.

Zasada działania rejestru przesuwającego, stosowanego w przetwornikach typu CCD

plikę ładunkową analizowanego obrazu. Do odczytu fotoładunków stosowane jest urządzenie elektroniczne przesuwające sekwencyjnie fotoładunki zgromadzone w światłoczułych elementach pamięci do wejścia przetwornika ładunek/prąd (q/U). Od tej charakterystycznej czynności, przetworniki wykorzystujące ten sposób adresowania noszą nazwę analizatorów obrazów z przesuwem ładunku lub inaczej: typu CTD (ang. *Charge Transfer Device*). Przesuw ładunków zachodzi w urządzeniach elektronicznych zwanych rejestrami przesuwającymi, najczęściej wykonanymi za pomocą technologii układów sprzężonych ładunkowo oznaczanej symbolem: CCD (ang. *Charge Coupled Device*), określa się je także mianem analizatorów CCD.

Matryca CCD została wynaleziona przez George Smith'a i Willard Boyle'a w 1969 roku. Matryca ta jest krzemowym układem scalonym, pokrytym siatką niewielkich elektrod nazywanych fotokomórkami, po jednej dla każdego piksela obrazu. W momencie kiedy światło dociera do fotokomórki, powoduje uwolnienie z niej pewnej liczby elektronów. Ponieważ każda komórka otoczona jest izolatorem, uwolnione elektrony pozostają w niej uwięzione. Fotokomórka gromadzi tym więcej elektronów, im więcej światła do niej dociera. Po naświetleniu matrycy CCD następuje pomiar napięcia wykonywany dla każdej komórki z osobna. W jego wyniku otrzymuje się informację o ilości elektronów zgromadzonych w danej komórce, a tym samym o ilości światła docierającego w to miejsce. Wynik pomiaru (próbkiwania) jest następnie przetwarzany w przetworniku analogowo-cyfrowym na postać cyfrową. Podczas naświetlania matrycy CCD fotoładunki są generowane i akumulowane w umieszczonych pionowo obok siebie analizatorach linii AL (rys. 22), nazywanych analizatorami kolumn. Rejestr adresujący każdego z analizatorów kolumn nie jest zakończony przetwornikiem q/U , lecz dołączony do przypisanego mu ogniwa rejestru przesuwającego CCD, zwanego rejestrze wyjściowym. Adresowanie pikseli w tej koncepcji polega na sekwencyjnym przesuwaniu zakumulowanej w analizatorach kolumn repliki fotoładunkowej do rejestru wyjściowego. Przesuw zachodzi skokowo co linię analizy obrazu w przedziale czasu przewidzianym na wygaszanie linii i polega na przesunięciu fotoładunków zgromadzonych w rejestrach analizatorów kolumn o jedno tylko ogniwo do przodu. Następnie ładunki w rejestrze wyjściowym są kolejno przesuwane do przetwornika q/U i odczytywane. Po odczytaniu wartości ładunków elektrycznych ze wszystkich komórek, matryca CCD jest gotowa do zarejestrowania kolejnego zdjęcia. Fotokomórki są elementami wrażliwymi wyłącznie na ilość światła, które pada na ich powierzchnię, i nie rozpoznają one na przykład koloru.

Jednoczesne wykorzystywanie komórek pamięci do przetwarzania i adresowania jest niestety przyczyną poważnej wady opisywanego wyżej rozwiązania. Podczas przesuwu fotoładunków do transferowanych pakietów dodają się obce fotoładunki aktualnie wytwarzane w ogniwach rejestru, przez co łączna ilość fotoładunków w transferowanych pakietach przestaje być proporcjonalna do natężenia oświetlenia piksela w ich pozycji startowej. Prowadzi to do istotnego zniekształcenia procesu przetwarzania, przejawiającego się m.in. zmianą jasności tła wzdłuż linii analizy (smużenie) oraz pojawianiem artefaktów (plamek) w analizowanym obrazie, stąd nosi ono nazwę zaplamienia. Zaplamienie można wyeliminować jedynie przez zaciemnienie (zastąpię-



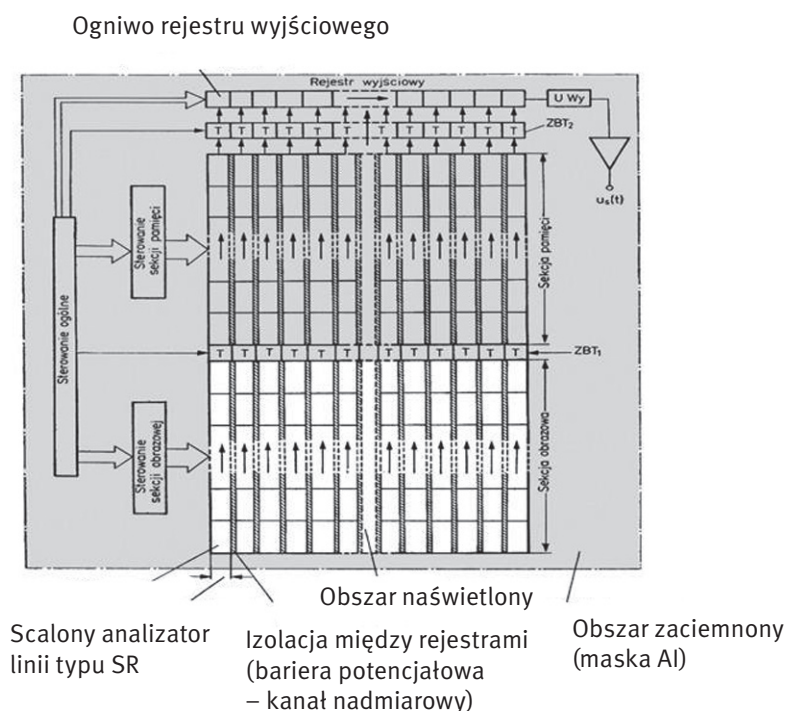
Rysunek 22.
Budowa i zasada działania matrycy CCD

cie) przetwornika podczas transferu ładunków, np. za pomocą przestony mechanicznej lub ciekłokrystalicznej. Oba sposoby są jednak dość powolne i nie zawsze możliwe do zastosowania. W takich przypadkach natężenie zaplamienia można jedynie zmniejszyć minimalizując czas transferu ładunków, a więc zwiększając częstotliwość przebiegu zegarowego, co podwyższa koszt przetwornika.

Wad tych jest pozbawiony analizator CCD z przesuwem półobrazu (rys. 23). Sekcja obrazowa jest naświetlana analizowanym obrazem świetlnym (obszar jasny na rysunku). Wyjścia wszystkich analizatorów kolumn są dołączone do identycznego zestawu rejestrów przesuwających CCD, umieszczonych w części zacienionej przetwornika (obszar zacieniony na rysunku), tworzących tzw. sekcję pamięci. Wyjścia rejestrów przesuwających sekcji pamięci dołączone są do odpowiadających im ogniów rejestru wyjściowego zakończonego przetwornikiem q/U. W pierwszym etapie obraz jest rejestrowany w sekcji obrazowej, następnie realizowany jest transfer ładunków do sekcji pamięciowej. Ponieważ na tym etapie nie ma konwersji q/U, transfer może być zrealizowany szybko, eliminując zaplamienie. W trzecim etapie zachodzi konwersja q/U, ale w sekcji zacienionej – może więc zachodzić wolno. Zastosowanie w opisywanym rozwiązaniu sekcji pamięci zminimalizowało efekt zaplamienia i pozwoliło na maksymalne wykorzystanie powierzchni światłoczułej w sekcji obrazowej, dzięki czemu analizatory CCD tego typu zapewniają obecnie najwyższą zdolność rozdzielczą spośród wszystkich rozwiązań analizatorów.

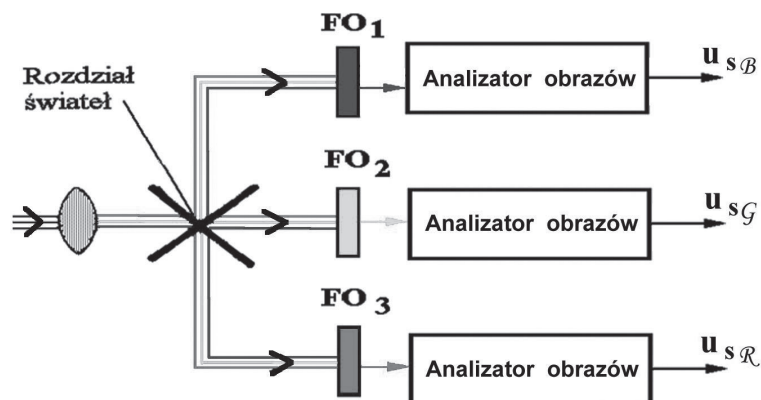
Istnieje jeszcze jedna szeroko stosowana klasa przetworników, a mianowicie **przetworniki typu CMOS APS** (ang. *Active Pixel Sensors*). Przetworniki te opracowano w drugiej połowie lat 90. XX wieku głównie w celu zmniejszenia kosztów produkcji sensorów i zmniejszenia poboru mocy. Przetworniki CMOS można łatwo zintegrować w układach scalonych. Możliwe jest nawet opracowanie kamery na jednym układzie scalonym. Takich możliwości, ze względów technologicznych, nie dają przetworniki CCD.

Matryca CMOS działa na tej samej zasadzie co matryca CCD. Światło padające na elementy światłoczułe generuje w nich ładunki elektryczne. A więc pojedynczy piksel w matrycach CCD i CMOS jest praktycznie taki sam. Dopiero otoczenie piksela jest w matrycy CMOS zupełnie inne. W odróżnieniu od matryc CCD, w matrycach CMOS każdy piksel ma swój przetwornik ładunku na napięcie, każdy piksel ma swój adres i jego zawartość może być odczytana w dowolnej kolejności. Współczynnik wypełnienia, czyli stosunek sumy powierzchni wszystkich pikseli do powierzchni całej matrycy, jest mniejszy dla matryc typu CMOS. Efektem tego jest nieco mniejsza czułość matryc CMOS. Część światła wpadająca przez obiektyw pada na elementy elektroniki wbudowane w matrycę i nie jest zamieniana na ładunki elektryczne. Stąd wczesne implementacje przetworników CMOS miały niską czułość i charakteryzowały się dużymi szumami. Obecnie matryce CMOS mają prawie takie same parametry jak CCD.



Rysunek 23.
Analizator CCD z przesuwem półobrazu

Zarówno przetworniki CCD jak i CMOS nie umożliwiają w sposób bezpośredni analizy obrazów barwnych. W przypadku jednoczesnej, trójchromatycznej analizy światła wytwarzane są trzy sygnały dla barw podstawowych czerwonej, zielonej i niebieskiej o wartościach proporcjonalnych do udziału tych barw w analizowanym świetle.



Rysunek 24.
Trójprzetwornikowa analiza obrazu barwnego

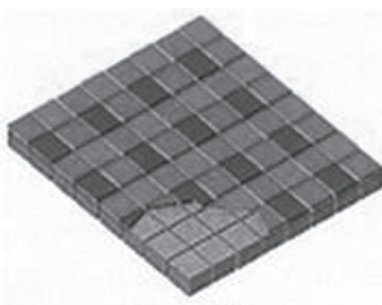
Klasycznym rozwiązaniem jednoczesnej, trójchromatycznej analizy obrazów świetlnych jest zastosowanie trzech analizatorów obrazów, które dokonują analizy tego samego obrazu kierowanego do każdego z analizatorów poprzez system rozdziálu światła (rys. 24). Taki sposób realizacji trójchromatycznej analizy obrazów świetlnych jest określany nazwą **analizy trójprzetwornikowej**.

Zastosowanie trzech analizatorów obrazów do analizy obrazów świetlnych jest rozwiązaniem kosztownym. Stąd w kamerach powszechnego użytku, dla których cena jest wyróżnikiem, stosuje się rozwiązanie uproszczone, w którym jednoczesną trójchromatyczną analizę obrazu barwnego wykonuje jeden, odpowiednio wy-



konany analizator obrazów. Realizacja trójdrożnej analizy światła przez jeden przetwornik jest możliwa, jeżeli powierzchnia jego materiału światłoczułego zostanie podzielona na wzajemnie rozłączne obszary, naświetlane tym samym obrazem, poprzez filtry.

Powierzchnia matrycy jest przykryta filtrem mozaikowym, który ogranicza działanie pojedynczego punktu światłoczułego na jeden z kolorów podstawowych. Z każdego punktu uzyskujemy więc dane o jasności fotografowanego motywu ale tylko dla jednej składowej koloru. W najbardziej popularnym układzie mozaika wygląda następująco: na jeden piksel czerwony lub niebieski przypadają dwa zielone (rys. 25). Odpowiada to warunkom widzenia człowieka, który najbardziej reaguje na zmiany jasności w zielonej części widma. Znając zasadę wykonania filtru i adresowania jego powierzchni światłoczułej można łatwo przypisać aktualnie przetwarzanej informacji elektrycznej barwę segmentu filtru, spod którego została pobrana i skierować ją do właściwego toru sygnału obrazu barwy podstawowej. Aby uzyskać dane o kolorze danego punktu musimy skorzystać z algorytmu interpolacji (przybliżania) danych z sąsiednich pikseli. Dopiero po procesie interpolacji uzyskujemy dane o kolorze danego punktu.



Rysunek 25.
Barwny filtr mozaikowy

Istnieją różne rodzaje filtrów mozaikowych. Filtry takie nazywane są **filtrami Bayera** od nazwiska ich wynalazcy. Przetwarzany przy ich pomocy sygnał wizyjny umożliwia zatem uzyskanie zielonego obrazu składowego z rozdzielczością niewiele ustępującą analizie trójprzetwornikowej. Jest to istotne, ponieważ w zakresie postrzegania odcieni zieleni oko ludzkie cechuje najwyższa ostrość widzenia zabarwienia drobnych elementów obrazu. Ta cecha filtrów Bayera została okupiona silnym spadkiem zdolności rozdzielczej analizy obrazów przetwarzanych spod segmentów R (czerwonych) i B (niebieskich) filtru, które zapełniają równomiernie i po równo pozostałą jego powierzchnię. Liczba każdego z tych segmentów jest równa zaledwie 1/4 maksymalnej liczby pikseli analizy, stąd zdolność rozdzielcza analizy w tych zakresach widma będzie 4-krotnie mniejsza od maksymalnej zdolności rozdzielczej przetwornika. Nie jest to jednak szczególna wada przetwornika, ponieważ zdolność oka do rozróżniania zabarwienia szczegółów w zakresie długo- i krótkofalowym (a więc czerwieni i błękitów) jest wielokrotnie niższa, stąd utrata informacji o zabarwieniu szczegółów w tych zakresach nie zostanie w ogóle dostrzeżona przez oko.

Zadanie 1. Pozyskiwanie materiału wideo

Sprzęt potrzebny do wykonania zadania:

- kamery wideo pracujące w standardzie mini dv: Panasonic NV-GS400 mini DV 3xLCD i Sony DCRTRV14E mini DV, kamery internetowe, telefon komórkowy z opcją nagrywania wideo,
- cyfrowy aparat fotograficzny.

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

Montaż wideo dzięki dynamicznemu rozwojowi technik cyfrowych, może być wykorzystywany zarówno w zastosowaniach profesjonalnych, jak i amatorskich. Słuchacze w trakcie trwania tego kursu zostaną zapoznani

z możliwościami pakietu do nieliniowej edycji materiału wizyjnego **Pinnacle Studio 9** oraz **Adobe Premiere Elements 7**. Oba programy stanowią kompletny zestaw narzędzi do cięcia, edycji, miksowania materiałów audio i wideo. Pakiety te umożliwiają dodawanie różnorodnych efektów specjalnych i innych elementów umożliwiających zaawansowaną edycję i montaż materiału wizyjnego

Pinnacle Studio to program komputerowy do domowej obróbki video, wyprodukowany przez Pinnacle Systems Inc. Udostępnia funkcje dodawania napisów, efektów dźwiękowych, komentarzy i efektów specjalnych. Obsługuje większość znanych formatów wraz z HD. Ma funkcję przechwytywania filmów wideo ze źródeł cyfrowych (format miniDV, Digital8 lub MicroDV) – kamer HDV, telefonów komórkowych, ze źródeł analogowych (format VHS, SVHS, VHS-C lub SVHS-C). Po zainstalowaniu dodatkowej karty możliwe są import i edycja materiału video bezpośrednio w formacie MPEG-1 lub MPEG-2. Program Pinnacle Studio umożliwia edycję poklatkową oraz zapis na płytach blue-ray.

Adobe Premiere Elements jest aplikacją do nieliniowej pracy nad materiałami wideo. Umożliwia montaż i obróbkę materiału audio-wideo, zarówno wysokiej, jak i niskiej rozdzielczości. Udostępnia typowe dla swojej klasy oprogramowania funkcje, takie jak: moduł do przechwytywania materiału wideo z kamer i magnetowidów, korektor koloru, mikser dźwięku surround, kluczowanie oraz zaawansowany moduł eksportu. Współpracuje z wieloma kartami wideo, które rozszerzają możliwości pracy programu w czasie rzeczywistym oraz ze sterownikami do kontroli funkcjami programu.

Mimo, że na rynku oprogramowania istnieje wiele programów dedykowanych do montażu nieliniowego, wszystkie one mają podobny charakter. Wyglądają one jak wirtualne studio montażowe, ze stołem montażowym, oknem z materiałami filmowymi oraz innymi elementami niezbędnymi w pracy montażysty. Cały proces jest wykonywany z wykorzystaniem technologii *drag and drop*, dzięki temu użytkownik w prosty sposób może wyciągać z wirtualnej biblioteki materiałów wizyjnych pliki wideo i układać je na wirtualnym stole montażowym. Pinnacle Studio 9 i Adobe Premiere Elements 7 oferują niemal wszystko to, co powinno być wykonywane przez program do cyfrowego montażu nieliniowego, czyli edycję obrazu wideo, obsługę kodów czasowych, synchronizację dźwięku i obrazu. W skład pakietów wchodzi również m.in. biblioteka efektów przejścia (tzw. tranzycje) oraz biblioteka filtrów obrazu wideo. W zasadzie wszystkie zintegrowane pakiety programowe do nieliniowego montażu wideo oferują te same opcje. Różnice mogą pojawić się w interfejsie komunikacji z użytkownikiem, tzn. poszczególne programy mogą być trochę inaczej obsługiwane. Z dwóch wymienionych pakietów Pinnacle Studio może się wydawać początkującemu użytkownikowi łatwiejszy do oprowadzania.

Cyfrowa edycja materiału wideo kojarzy się zwykle z drogimi i skomplikowanymi aplikacjami, wymagającymi potężnych komputerów. Warto tu wspomnieć, że istnieje edytor filmów **Windows Movie Maker**. Jest to bardzo prosty edytor o mocno ograniczonych możliwościach, ale to jest program typu freeware, można więc za darmo pobrać go z Internetu. Po raz pierwszy został dołączony do systemu operacyjnego Windows ME. W wielu przypadkach, opcje, jakich dostarcza edytor, są wystarczające. Należą do nich: dzielenie filmu na klipy, dodawanie komentarza audio lub łączenie obrazów tak, aby były w jednym ciągu. Aplikacja umożliwia szybkie zmontowanie cyfrowego filmu, udźwiękowienie go i zapisanie na płycie DVD, taśmie wideo lub dysku twardym komputera. Pliki zapisywane są w formacie Windows Media Video (wmv).

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 min., 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). W kamerach amatorskich miniDV stosuje się kasety miniDV, które są niewiele większe o pudełka zapatek. Obraz jest zapisywany na taśmie o szerokości 6,35 mm. W przypadku kaset Digital 8 szerokość wynosi 8 mm.



Na wstępie należy zapoznać się ze sprzętem, zwłaszcza z obsługą i funkcjami kamer cyfrowych. Szczególnie ciekawa jest kamera Panasonic NV-GS400 mini DV 3xLCD. Na uwagę zasługuje system 3xCCD, czyli potrójna matryca CCD, umożliwiająca na 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. To wszystko dzięki technologii miniaturyzacji przetwornika. Często, specjalnie dla systemu 3CCD specjalizowane procesory (w przypadku kamery Panasonic – tzw. *Crystal Engine*) przetwarzają sygnał niezależnie dla każdej matrycy. Jedną z funkcji wykorzystujących te możliwości jest redukcja szumów. Wydobywa szczegóły z każdej barwy RGB nawet przy słabym oświetleniu. Nagrywanie filmów i robienie zdjęć ułatwia szeroko otwierany ekran LCD. Szybki wybór trybu pracy umożliwia pokrętko umieszczone z tyłu kamery. Wystarczy wybrać i gotowe. Innym udogodnieniem jest uchwyt, którego wymodelowany i dopasowany do dłoni kształt zwiększa stabilność kamery i zmniejsza zmęczenie ręki. Kamera może pracować w trybie filmowym Pro-Cinema. Umożliwia to kręcenie filmów z prędkością prawie 25 kl/s, czyli porównywalną do prędkości stosowanej w filmach kinowych (24 kl/s). Dzięki temu można odtwarzać obrazy na szerokokątnym ekranie telewizora ze wzmocnionymi basami i wyciszeniem wiatru. Poza tym filmy są wyświetlane w płynny sposób, tak jak na ekranie w kinie. Za pomocą wejścia DV można podłączyć sprzęt do komputera i wcześniej poddany edycji film z powrotem skopiować na kamerę. Natomiast wejście analogowe i S-Video (zapewnia wyższą jakość obrazu w systemie DV) umożliwiają nagrywanie obrazu z magnetowidu lub telewizora. Zresztą w ten sam sposób można wykorzystać wyjścia z kamery i nagrać materiał na komputer, DVD czy magnetowid. Kamera jest wyposażona w strumieniową transmisję USB, dzięki której można przysyłać filmy w formacie MPEG4 lub zdjęcia bezpośrednio przez Internet. Poza tym umożliwia funkcję kamery internetowej.

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),
- muzykę mogącą służyć za podkład do scen w filmie.

2 PRZETWARZANIE SEKWENCJI WIDEO

2.1 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 przy pomocy 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 nagrać zmontowany materiał wideo 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 zniekształcenia. Warto jednak zauważyć, że pasmo sygnału nieskompresowanego: $720 \times 576 \times 25$ daje ok. 20 MB/s nagrania (kolor jest zapisany na 16 bitach), co jest w większości przypadków nie do przyjęcia. Kompresja DV ma ustalony na stałe współczynnik 5:1, a więc redukuje pasmo do 3,7 MB/sek. na obraz, do czego dochodzi $2 \times 16 \times 48000 = 1,5$ MB/sek. na dźwięk, dając nam łącznie ok. 5 MB/sek., czyli znacznie redukuje ilość danych.

W przypadku kart analogowych proces przechwytywania materiału wideo często jest nazywany **digitalizacją**. W trakcie przeprowadzania tego procesu analogowy materiał wideo zostaje poddany złożonemu procesowi, podczas którego jest zamieniany 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 i 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.

2.2 STANDARDY KODOWANIA SEKWENCJI WIDEO MPEG

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-video 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 zapisania takiego obrazu cyfrowo powoduje, że każda sekunda obrazu zajmuje około 30 megabajtów. Na początku lat 90. XX wieku został opracowany standard kompre-



sji **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, został przystosowany 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 jest stosowana 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*), takie obszary 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 przesuujące się obiekty i prawidłowo zostanie opisany ich ruch, metoda ta jest nazywana 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 **luminacji** (jasności) oraz dwie składowe **chrominancji** (koloru) – YCRCB. Poszczególne składowe są obliczane na podstawie wartości dla poszczególnych kolorów RGB mnożonych przez różne współczynniki:

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

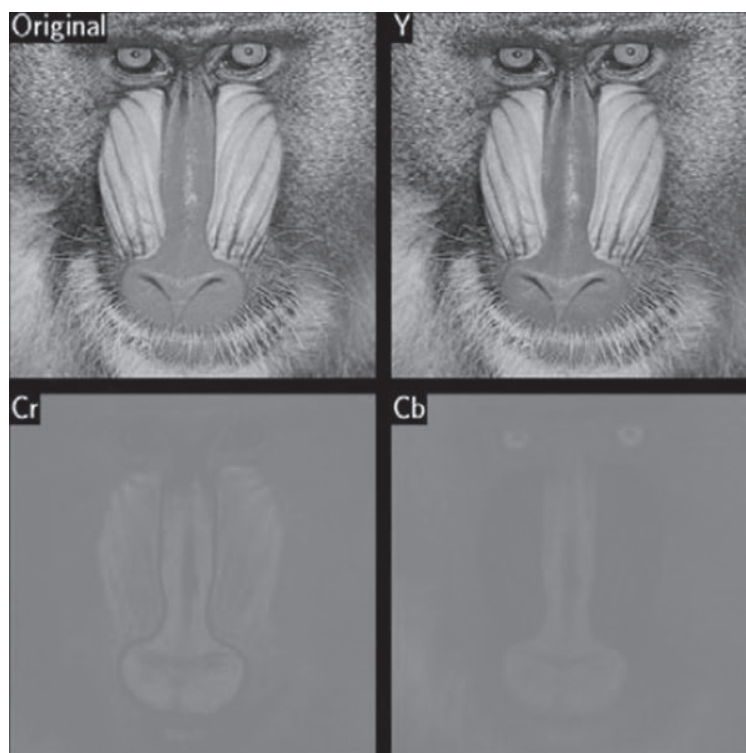
$$C_b = 0.564(B - Y),$$

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

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

Specyfika ludzkiego wzroku, dla którego bardziej znacząca 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. 26.

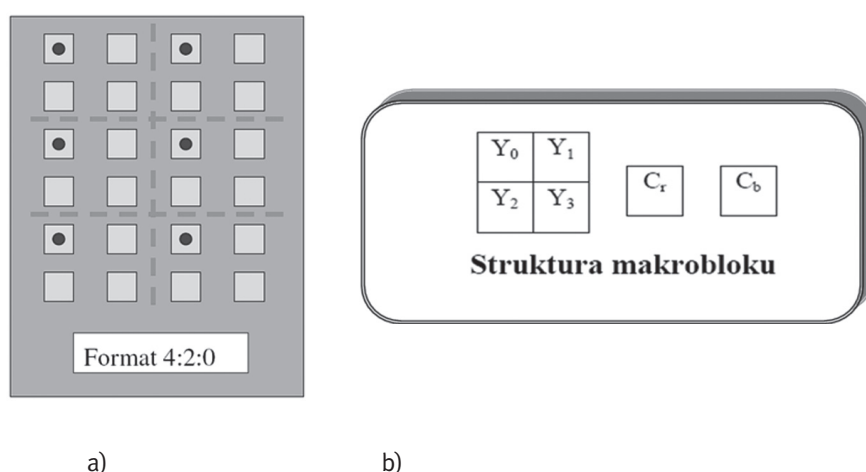




Rysunek 26.

Składowe luminancji Y i chrominancji CR CB obrazu kolorowego

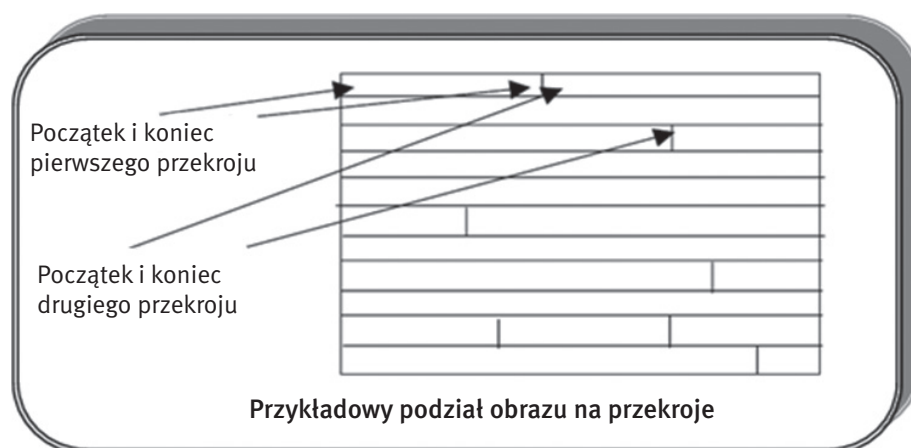
Dla standardu MPEG-1 przyjęto strukturę próbkowania obrazu 4:2:0 (na 4 próbki luminancji przypadają dwie próbki chrominancji w jednej linii, oraz 0 próbek chrominancji w kolejnej linii). Taka struktura próbkowania jest przedstawiona na rys. 27a. Wartości brakujących próbek chrominancji wyznacza się poprzez interpolację (obliczenie wartości pośrednich). Elementarną porcję obrazu kodowaną przez koder MPEG nazywa się **makroblokiem**. Makroblok to tablica 16x16 pikseli złożona z próbek luminancji i odpowiadające im tablice chrominancji CR i CB o rozmiarze 8x8 pikseli (rys. 27b).



Rysunek 27.

Format próbkowania obrazu 4:2:0 stosowany w standardzie MPEG-1 oraz struktura makrobloku

W następnym kroku makrobloki są składane w tzw. **przekroje**, które mogą zawierać różną ilość makrobloków (rys. 28). Niezależne kompresowanie każdego przekroju zabezpiecza przed rozszerzaniem się ewentualnych błędów na cały obraz.



Rysunek 28.

Podział obrazu na przekroje

W dalszych etapach kodowania MPEG występuje kilka procesów, które wymagają dodatkowego wyjaśnienia. Należą do nich dyskretna transformacja kosinusowa, kwantyzacja i kodowanie Huffmana.

Dyskretna transformacja kosinusowa (DCT) pomaga rozdzielić sygnał na części (lub pasma spektralne), przekształcając dane do postaci umożliwiającej zastosowanie efektywnych metod kompresji. Transformacja DCT nie służy do kodowania a jedynie do przekształcenia danych odpowiadających wartościom amplitud pikseli danego bloku 8x8 pikseli do postaci umożliwiającej efektywne zastosowanie metod kompresji. Transformacja DCT przetwarza obraz z obszaru przestrzennego do częstotliwościowego. W wyniku działania transformaty na sygnale wejściowym powstają odpowiadające mu współczynniki transformaty. Należy podkreślić, że wszystkie mnożenia podczas transformaty DCT operują w przestrzeni liczb rzeczywistych. 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. 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. Proces kwantyzacji można opisać równaniem:

$$k(x) = \text{round} \left(\frac{F(x)}{Q(x)} \right)$$

gdzie $F(x)$ jest współczynnikiem transformacji, $Q(x)$ jest tablicą kwantyzacji, a $\text{round}(x)$ jest funkcją zaokrąglającą x do najbliższej liczby całkowitej.

Tablica kwantyzacji odpowiada za stopień kompresji. Kwantyzacja jest procesem kompresji stratnej. Informacja utracona w wyniku kwantyzacji nie może być odzyskana. Im większy stopień kompresji tym gorsza jakość odtworzonego obrazu. Tablice kwantyzacji są dobierane na podstawie eksperymentów i analizy między innymi zdolności percepcyjnej oka ludzkiego osobno dla luminancji, osobno dla składowych koloru.

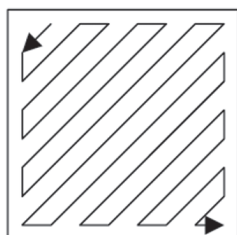
Kodowanie Huffmana jest to bezstratna metoda kodowania statystycznego, przedstawiona przez Davida Huffmana w roku 1952. Kodowanie Hoffmana stanowi jedną z najprostszych i łatwych w implementacji metod kompresji bezstratnej. W algorytmie wykorzystywany jest fakt, że pewne wartości danych występują częściej niż inne. Jeżeli zatem zakodujemy częściej występujące wartości za pomocą krótszych słów kodowych, a rzadziej występujące za pomocą dłuższych to średnio sumaryczna długość zakodowanych danych będzie krótsza niż przed kodowaniem. Oznacza to konieczność wprowadzenia jednoznacznego systemu prefiksów – krótsze słowo kodowe nie może być początkiem dłuższego słowa kodowego (np. 110 i 1101001 nie mogą być jednocześnie słowami kodowymi).

2.2.1 PRZEBIEG PROCESU KODOWANIA W STANDARDACH MPEG-1 – 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 ramkach 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. 29). Wybieranie zygzakowe jednocześnie szereguje współczynniki od najbardziej istotnych do najmniej ważnych z punktu widzenia kompresji obrazu.

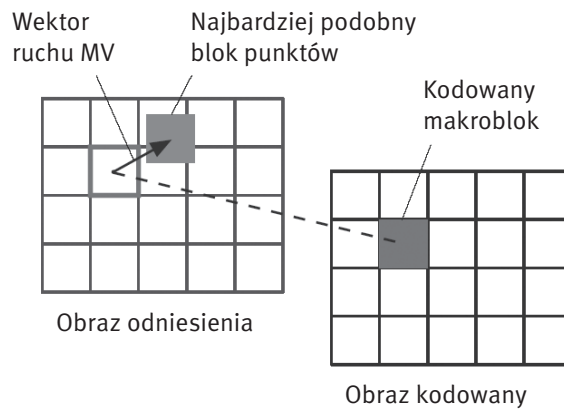


Rysunek 29.

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 jest tworzona 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.

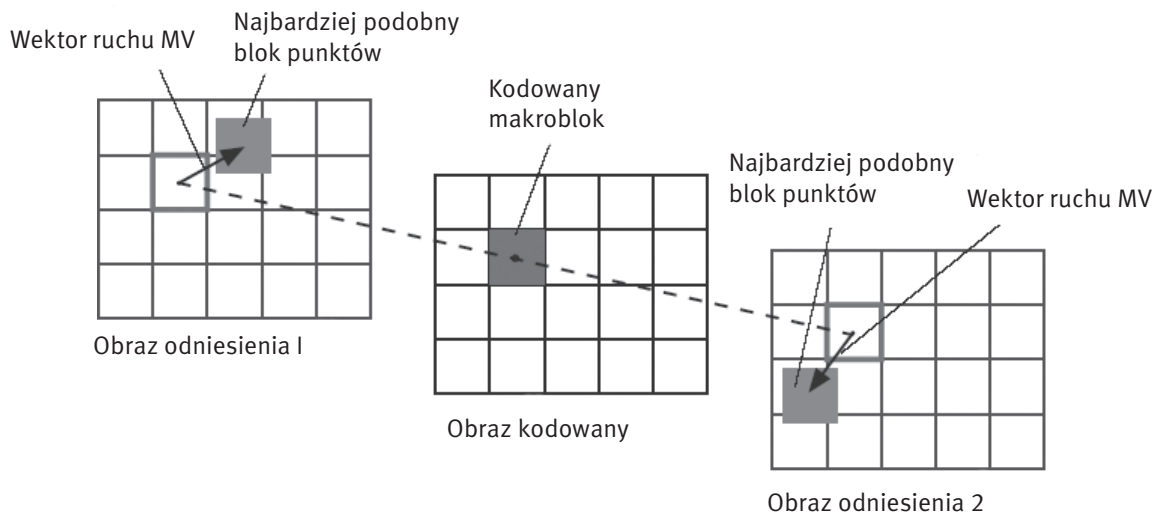


Rysunek 30.

Sposób kodowania obrazów typu P

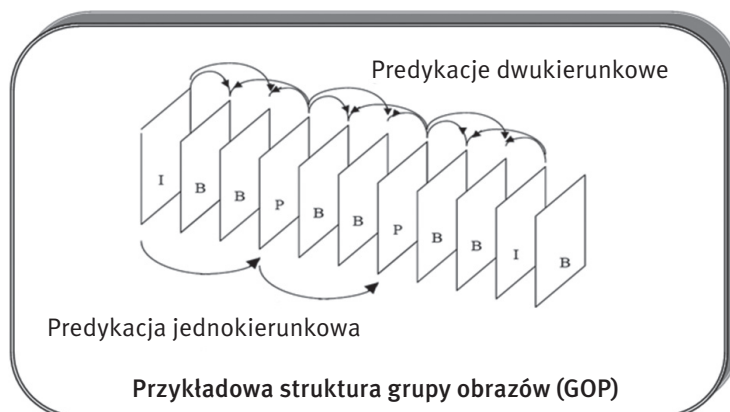
Dwa kolejne obrazy są odejmowane od siebie i ich różnicę koduje się tak, jak omówione wcześniej obrazy typu I, wykorzystując transformatę kosinusową i kwantyzację jej współczynników. Ponieważ jest to kompresja stratna, więc transmitowana informacja różni się od wyniku odejmowania kolejnych obrazów. Ta 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 dekoderze – obraz ten jest nazywany **obrazem odniesienia** (rys. 30). 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.

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 jest tworzona przez uśrednienie informacji pochodzących od obu obrazów odniesienia, dlatego ten typ obrazu jest nazywany **obrazem interpolowanym** (rys. 31). Metoda kodowania obrazów typu B wymaga, aby dekoder wcześniej znał oba obrazy odniesienia. Obrazy odniesienia muszą być wysłane wcześniej niż zakodowane obrazy typu B. Bez nich niemożliwe jest rozkodowanie.



Rysunek 31.

Sposób kodowania obrazów typu B



Rysunek 32.

Struktura grupy obrazów

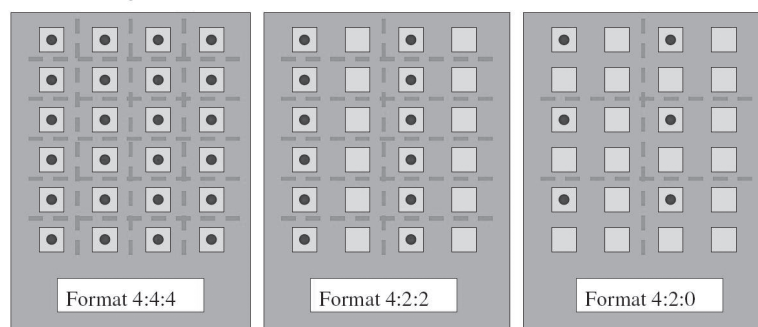
Poszczególne typy obrazów I, P i B stosowane w standardzie MPEG zapewniają różne stopnie kompresji. Najmniejszy stopień kompresji umożliwiają obrazy typu I, gdyż zawierają pełną informację o obrazie. Obrazy te muszą być jednakże dość często transmitowane, gdyż odbiornik może rozpocząć dekodowanie sekwencji 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 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 treści przekazywanego programu. Przykładowa struktura grupy obrazów jest przedstawiona na rys. 32.

2.2.2 RÓŻNICE POMIĘDZY STANDARDAMI MPEG-1 I MPEG-2

Standard MPEG-2 został opracowany z myślą o zastosowaniu w cyfrowych systemach telewizyjnych, do obsługi których nie nadawał się MPEG-1. 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/50 Hz, jak i amerykańskim 525 linii/60 Hz. 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 jest przesyłana 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.

W standardzie MPEG-2 jest wykorzystywana struktura próbkowania 4:2:0 lub 4:2:2 (rys. 33). W strukturze próbkowania 4:2:0 następuje zmniejszenie rozdzielczości sygnałów chrominancji w obu kierunkach: pionowym i poziomym.

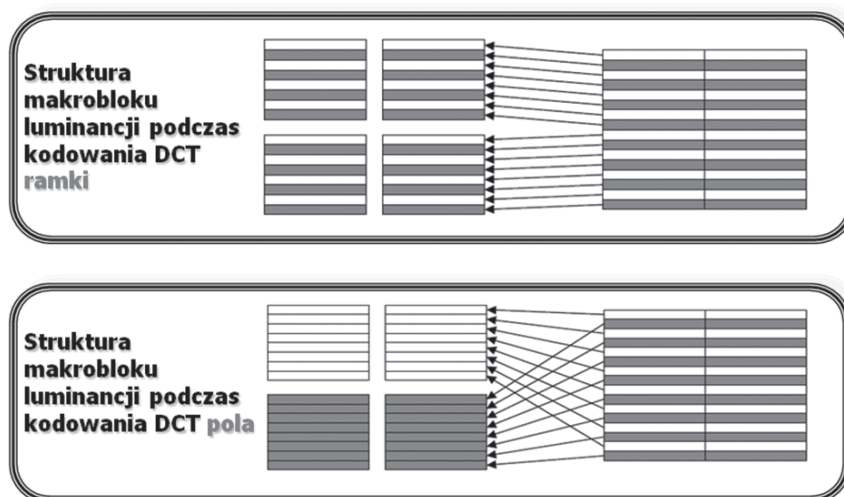


Rysunek 33.

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

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 .

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. 34.



Rysunek 34.

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 zakłócona i lepiej jest użyć kompensacji ruchu.

Standard MPEG-2 może być wykorzystywany dla kodowania obrazów o różnej rozdzielczości przy zastosowaniu różnych wariantów kompresji sygnałów. W tym celu przyjęto dwa podstawowe pojęcia poziom (ang. *level*) oraz profil (ang. *profile*). Pojęcie poziom jest związane z rozdzielczością kodowanego obrazu. Dla każdego z poziomów można stosować różne warianty metody kompresji sygnału, pozwalające na uzyskanie różnych

prędkości bitowych. Parametry te nazwano profilem. Najważniejsze ze stosowanych w systemie MPEG-2 profile to (rys. 35):

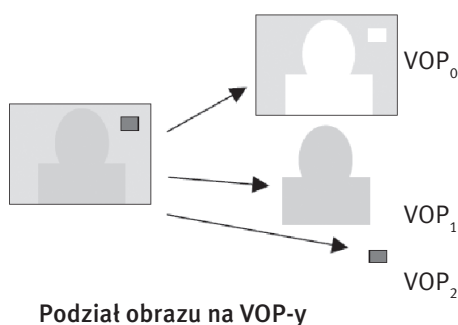
- Profil prosty (ang. *Simple Profile*) wykorzystuje dyskretną transformację kosinusową, kodowanie z prognozowaniem (obrazy typu P) oraz kompensację ruchu, przy czym sygnały różnicowe kolorowości obrazu są kodowane sekwencyjnie, co drugą linię (standard 4: 2: 0).
- Profil główny (ang. *Main Profile*) wykorzystuje tę samą metodę kodowania, lecz dwa rodzaje prognozowania (ramka P i B), co zapewnia lepszą jakość sygnału przy tej samej prędkości bitowej.
- Profil skalowany szumowo (ang. *SNR Scaleable Profile*). Dane wizyjne są podzielone na części: sygnał podstawowy i sygnał podwyższający jakość. Sygnał podstawowy wytwarza obraz o zmniejszonym stosunku sygnału do szumu i wymaga znacznie mniejszej prędkości bitowej, może więc być przesyłany przez tory transmisyjne o mniejszej przepustowości. Dodanie do sygnału podstawowego sygnału podwyższającego jakość poprawia stosunek sygnału do szumu w obrazie. Zapewnia to przy tej samej prędkości bitowej lepszą jakość odtwarzanego obrazu niż w profilu głównym.
- Profil skalowany przestrzennie (ang. *Spatially Scaleable Profile*). Jest zbudowany podobnie jak profil skalowany szumowo, lecz wykorzystuje inną metodę podziału danych, a mianowicie w zależności od rozdzielczości obrazu. Sygnał podstawowy odpowiada obrazowi o zmniejszonej rozdzielczości. Obraz pełnej rozdzielczości otrzymuje się przez zsumowanie tego sygnału z sygnałem poprawiającym jakość.

	Profil prosty SIMPLE	Profil główny MAIN	Profil skalowalny jakościowo SNR	Profil skalowalny przestrzennie SPATIAL	Profil wysoki HIGH	Profil 4:2:2 PROFILE 4:2:2
Poziom wysoki HIGH		4:2:0 1920x1152 80 Mb/s I, P, B			4:2:0, 4:2:2 1920x1152 100 Mb/s I, P, B	
Poziom wysoki 1440 HIGH1440		4:2:0 1440x1152 60 Mb/s I, P, B		4:2:0 1440x1152 60 Mb/s I, P, B	4:2:0, 4:2:2 1440x1152 80 Mb/s I, P, B	
Poziom główny MAIN	4:2:0 720x576 15 Mb/s I, P	4:2:0 720x576 15 Mb/s I, P, B	4:2:0 720x576 15 Mb/s I, P, B		4:2:0, 4:2:2 720x576 20 Mb/s I, P, B	4:2:2 720x608 50 Mb/s I, P, B
Poziom niski LOW		4:2:0 352x288 4 Mb/s I, P, B	4:2:0 352x288 4 Mb/s I, P, B			

Rysunek 35.
Poziomy i profile w standardzie MPEG-2

2.2.3 KOMPRESJA MPEG-4

Kompresja MPEG-4 stawiała za cel osiągnięcie wysokiej jakości przy bardzo niskich przepływnościach strumienia danych 10kb/s-1Mb/s. To, co odróżniało MPEG-4 od pozostałych standardów z tej grupy, to podejście obiektowe. Każda scena do zakodowania jest przedstawiana w oparciu o obiekty audiowizualne AVO (rys. 36).



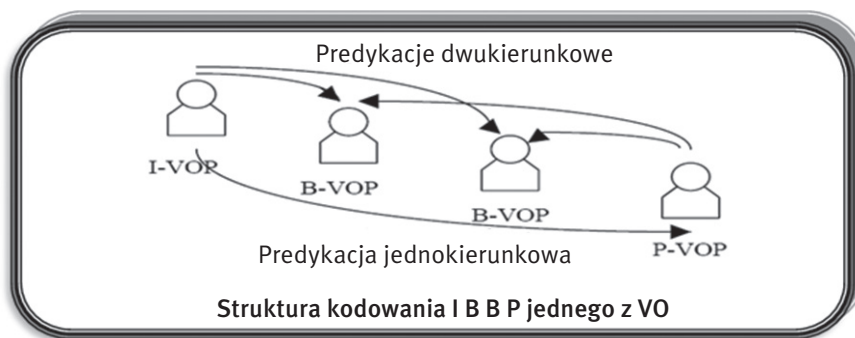
Rysunek 36.
Podział obrazu na obiekty wideo

Koncepcja standardu MPEG-4 zakłada przejście od pojedynczego algorytmu kompresji sekwencji obrazów, do zbioru narzędzi kodowania, które łączone ze sobą w odpowiedni sposób tworzą algorytmy kompresji przeznaczone do poszczególnych zastosowań. MPEG-4 może wykorzystywać osobne narzędzia do kodowania kształtu, tekstury, ruchu, skalowalności, syntetycznych obrazów i wiele innych.

Wejściowa sekwencja video jest zbiorem kolejnych obrazów oddzielonych w czasie. Koncepcja obiektów video (VO) oraz ich próbek czasowych (VOP) jest tu kluczowa. VOP może być w pełni określony przez jego **kształt i teksturę**. W scenach naturalnych VOP-y podlegają półautomatycznej lub automatycznej **segmentacji**, a informacja o wypadkowych kształcie jest reprezentowana przez binarną maskę kształtu lub maskę skali szarości. VO mogą podlegać podziałowi na **warstwy (VOL)**, umożliwiające skalowalną reprezentację obiektów video. Warstwa podstawowa zapewnia bazową rozdzielczość czasową i przestrzenną, pozostałe warstwy ulepszające, kodowane w odniesieniu do warstwy podstawowej, pozwalają uzyskać lepszą rozdzielczość czasową lub przestrzenną. Każdy VO jest kodowany niezależnie. Wraz z VO koder przesyła informacje o kompozycji sceny, opisujące gdzie i kiedy każdy VO ma być wyświetlany. Ze względu na wydajność oraz wsteczną kompatybilność, VO kompresowane są przez kodowanie odpowiadających im VOP-ów, z wykorzystaniem makrobloków 16x16 pikseli.

Koder VO musi zawierać tor kodowania kształtu i tor kodowania tekstury. W aplikacjach uproszczonych, w których VOP jest prostokątną ramką, tor kodowania kształtu może być zablokowany. Pochodząca z MPEG-2 koncepcja kodowania obrazów zostaje rozszerzona na kodowanie VOP-ów: intra (I-VOP), predictive (P-VOP), i bidirectionally predictive (B-VOP) (rys.37). MPEG-4 wykorzystuje kodowanie z kompensacją ruchu i transformacją DCT, jednak zarówno estymacja ruchu jak i kompensacja ruchu musi uwzględniać dowolność kształtu VOP.

Kodowanie kształtu jest realizowane często przy użyciu **maski binarnej**. Informacja o masce binarnej jest reprezentowana przez macierz (bitmapa) o rozmiarach takich, jak prostokątna ramka ograniczająca VOP (w pionie i poziomie musi być wielokrotnością 16 pikseli). Pikselowi wewnątrz ramki ale na zewnątrz VOP jest przypisywana wartość 0, pozostałym – 255. Ramka dzielona jest na bloki 16x16, te zawierające wszystkie zera są oznaczone jako ALL-0 i dla nich nie jest kodowana informacja o kształcie ani teksturze. Dla bloków zawierających wszystkie wartości = 255 tzw. ALL-255 konieczne jest kodowanie tekstury. Informacja o kształcie dla pozostałych bloków jest kodowana stratnie lub bezstratnie przy użyciu techniki z kompensacją ruchu. Możliwe jest również kodowanie kształtu przy użyciu **maski skali szarości**. W tym przypadku każdy piksel może przyjmować wartość z zakresu 0-255 odpowiadającą stopniowi przezroczystości piksela. Informacje o kształcie są kodowane tylko stratnie z użyciem transformacji kosinusowej DCT z kompensacją ruchu.



Rysunek 37.
Struktura kodowania obiektu wideo

Do kodowanie tekstur jest używana transformacja falkowa lub DCT adaptowana dla potrzeb kodowania VOP o dowolnych kształcie. Tekstura VOP jest dzielona na makrobloki 16x16. Bloki leżące na granicach VOP-ów mogą zawierać piksele nie należące do nich, zatem są kodowane nieco inaczej. Kodowanie bloków wewnętrznych: najczęściej kodowane są konwencjonalnie z użyciem transformacji DCT dla bloków 8x8. Bloki luminancji i chrominancji kodowane są oddzielnie. Kodowanie bloków granicznych jest trudniejsze, gdyż w tym przypadku uwzględniany musi być zmienny kształt tych bloków.

Kod źródłowy jednej z implementacji MPEG-4, utworzonej przez Microsoft do użytku w systemie Windows Media (pliki WMV), dostał się w ręce hakerów i został udostępniony w Internecie. Stworzone przez hakerów nielegalne oprogramowanie kodujące (kodek), rozpowszechnione w Internecie jako DivX ;-), szybko zostało zaadaptowane do użycia w znanych systemach operacyjnych, takich jak Windows, MacOS, Linux i BeOS. Pierwotny kodek DivX ;-), udostępniony przez hakerów, był oznaczony numerem wersji 3.11. Wkrótce potem na rynku pojawił się komercyjny program o nazwie DivX 4.0 (już bez znaczka ,;-)') – nie zawierał on już wykradzonego kodu, lecz był implementacją MPEG-4 napisaną od podstaw przez firmę DivX Networks (obecnie DivX Inc.). DivX jest obecnie zastrzeżonym znakiem towarowym tej firmy, natomiast samo oprogramowanie miało kilka kolejnych wersji, wprowadzających bardziej zaawansowane techniki kodowania.

W związku z faktem, iż oprogramowanie DivX® jest komercyjne, zamkniętym produktem firmy DivX Networks, społeczność wolnego oprogramowania stworzyła alternatywny kodek XviD, rozpowszechniany na licencji GNU GPL. Wykorzystuje on szerszy zakres możliwości standardu MPEG-4 niż DivX i daje możliwość bardziej precyzyjnej kontroli sposobu kompresji, dzięki czemu filmy kompresowane tym kodekiem często mają wyższą jakość, niż komercyjnym DivX. Formaty DivX/XviD jakością niewiele ustępują DVD, charakteryzuje je duża rozdzielczość i duża kompresja obrazu (przeciętny film pełnometrażowy można zmieścić w objętości 700 MB, czyli jednej płyty CD), jednak z momentem upowszechniania się ekranów LCD o dużej przekątnej, jak i telewizorów HDTV, jakość tego formatu pozostawia obecnie bardzo wiele do życzenia.

Zadanie 2. Zgrywanie sekwencji wizyjnej

Sprzęt potrzebny do wykonania zadania:

- komputer,
- kamery wideo pracujące w standardzie mini dv: Panasonic NV-GS400 mini DV 3xLCD i Sony DCRTRV14E mini DV, kamery internetowe, telefon komórkowy z opcją nagrywania wideo,
- cyfrowy aparat fotograficzny,
- oprogramowanie do montażu nieliniowego Pinnacle Studio 9, Adobe Premiere Elements 7.

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

W przypadku kart analogowych, proces przechwytywania materiału wideo jest często nazywany **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 kasyety na dysk. Warto dodać, że kamera cyfrowa już w trakcie filmowania, aby zmniejszyć strumień danych, kompresuje obraz w stopniu 5:1 (kompresja DV). Przyjętym standardem jest DV. Ze względu na zastosowaną kompresję w trakcie przechwytywania nie można regulować jakichkolwiek parametrów zarówno obrazu, jak i dźwięku. Niezaprzeczalną 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.

Materiał wideo możemy przechwytywać jako duże partie materiału lub używając listy cięć (ang. *batch capture*) wybranych przez nas ujęć. Rozwinięciem tej metody jest tzw. skanowanie taśmy (ang. *scan tape*), 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 pakiety Pinnacle Studio 9 i Adobe Premiere Elements 7.

Pakiety Pinnacle Studio 9 i Adobe Premiere Elements 7 umożliwiają zgrywanie materiału filmowego zarówno bezpośrednio z cyfrowej kamery DV, jak i ze źródła analogowego za pomocą analogowej karty przechwytyjącej. 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.

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 i Adobe Premiere Elements 7,
- 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 EDYCJA I MONTAŻ OBRAZÓW WIDEO

3.1 TELEWIZJA ANALOGOWA I CYFROWA

Prace nad systemem telewizji kolorowej rozpoczęły się w połowie lat 50. XX wieku w Stanach Zjednoczonych. Nowy system musiał spełniać następujące założenia:

- nie mógł znacząco komplikować budowy odbiorników telewizji kolorowej, co mogło by wpływać na koszt produkcji odbiornika telewizyjnego i zmniejszyć jego dostępność dla widza ze względu na cenę;
- należało przyjąć zasadę odpowiedniości, czyli możliwości odbioru programu telewizji nadawanego w kolorze na odbiornikach czarnobiałych i odwrotnie;
- powinno być możliwe wykorzystywanie dotychczasowych kanałów częstotliwości do przesyłania sygnałów telewizji kolorowej, nie powodując zakłóceń w kanałach sąsiednich;
- jakość przesyłanego sygnału powinna być wysoka i zaspakajać wymagania widza.

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 składowe: luminancję Y i dwa sygnały różnicowe koloru Cr oraz Cb, czyli tzw. 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 próbki chrominancji na co drugiej linii. Takie ograniczenie rozdzielczości w pionie nie wpływa zasadniczo na jakość obrazu kolorowego, natomiast istotnie redukuje strumień informacji o obrazie.

W systemie NTSC obraz jest składany z 525 linii na ramkę, przy częstotliwości odświeżania 59,94 Hz (jest to skutkiem stosowania w USA częstotliwości prądu przemiennego wynoszącej 60 Hz) i 29,97 ramkach na sekundę. 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 50 Hz. 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.

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

3.1.2 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.

3.2 POPRAWA JAKOŚCI OBRAZU

Technika cyfrowa umożliwia zastosowanie wielu metod poprawy jakości obrazu przekazów telewizyjnych. Najczęściej spotykane zniekształcenia wynikają z pojawienia się artefaktów procesu kompresji. Do innych



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 trójwymiarowych filtrów cyfrowych, filtrów grzebieniowych, układów korekcji podstawy czasu i stosowaniu odpowiednich technik (100 Hz, obraz bez przepłotu). Poprawie jakości sprzyja też sztuczne podnoszenie rozdzielczości w oparciu o technikę nadpróbkowywania i interpolacji wartości pikseli.

Eliminacja migotania – technika 100 Hz – polega na podwajaniu częstotliwości powtarzania półobrazów. Wprowadzenie w standardzie PAL wybierania międzyliniowego z częstotliwością powtarzania półobrazów (pola) 50 Hz miało w założeniu doprowadzić do zmniejszenia efektu migotania jasnych płaszczyzn na ekranie telewizora. Zjawisko migotania obrazu staje się szczególnie dokuczliwe przy przekątnych większych niż 29 cali. W odbiornikach stosuje się więc podwajanie częstotliwości powtarzania półobrazów, czyli technikę 100 Hz. Może być ona realizowana w różnych wariantach. Załóżmy, że mamy sekwencję wizyjną złożoną z półobrazów (pól) A i B wyświetlanych co 20ms, składających się na całkowity obraz o rozdzielczości pionowej 576 linii. W wariancie AABB półobraz A zostaje wyświetlony dwa razy pod rząd co 10ms, a następnie tak samo reprodukowany jest półobraz B. Takie rozwiązanie eliminuje migotanie dużych jasnych powierzchni ekranu, ale wprowadza często bardziej dokuczliwe zjawisko migotania linii i konturów w obrazie. Tej wady nie ma sposób odtwarzania ABAB, wymagający jednak większej pamięci, zdolnej przechować dwa półobrazy. Wariant ten powoduje jednak zniekształcenia w odtwarzaniu szybko poruszających się obiektów (efekt rozdławiania). Obecnie stosuje się interpolację treści półobrazów, polegającą na utworzeniu na podstawie przesyłanej informacji nowych półobrazów A' i B'. Algorytmy interpolacyjne tak wyliczają wartości nowych pikseli, aby w rezultacie doprowadzić do poprawnego odtwarzania ruchu przy niezauważalnym migotaniu. Treść wizyjna wyświetlana jest z częstotliwością 100 Hz w kolejności AA'BB'.

3.2.1 REDUKCJA ARTEFAKTÓW WYNIKAJĄCYCH Z KOMPRESJI

Stosowanie stratnej kompresji wprowadza do rekonstruowanego obrazu wiele zniekształceń zwanych **artefaktami**. Mogą one powodować wrażenie istotnego pogorszenia jakości. Za powstanie artefaktów odpowiada zwykle koder MPEG-2 stosowany po stronie nadawczej. Do typowych zjawisk należy tutaj efekt blokowy. Jest on charakterystyczny dla metod kompresji bazujących na przetwarzaniu bloków pikseli. W procesie kwantyzacji składowe stałe reprezentujące sąsiednie bloki mogą być zakodowane z różną precyzją, co powoduje później widoczne różnice w luminancji tła fragmentów obrazu. Na ekranie pojawia się wtedy charakterystyczna struktura siatki. Innym artefaktem jest *mosquito noise*. Nazwa bierze się z faktu, że przypomina on chmarę komarów unoszących się nad obiektem w rytmie jego ruchów. Zjawisko wynika z tego, że w procesie kodowania MPEG-2 fragmenty obrazów odpowiadające wyższym częstotliwościom przestrzennym, a więc opisujące występujące w obrazie krawędzie, są kodowane z małą precyzją. Wspomniane artefakty są usuwane w układach filtrów cyfrowych. Zastosowanie prostych filtrów cyfrowych może jednak prowadzić do zmniejszenia wyrazistości obrazu lub innych efektów pogarszających jego subiektywną ocenę.

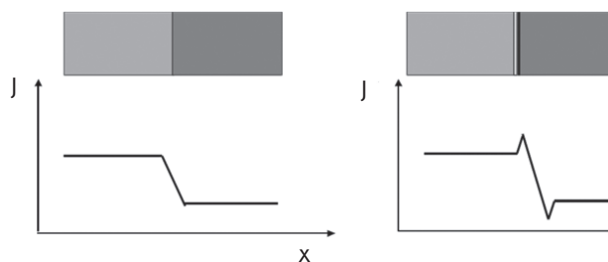
3.2.2 EKSPONOWANIE KONTURÓW OBRAZU

Poprawa ostrości konturów odbierana jest przez większość osób jako zwiększenie rozdzielczości. Już samo zwiększenie kontrastu, które powoduje większe różnice w jasności sąsiadujących fragmentów obrazu, potęguje wrażenie lepszej ostrości konturów. Jednak zwiększanie kontrastu w skali całego obrazu prowadzi do zatarcia się poziomów jasności w ciemnych i jasnych partiach obrazu. Stosuje się więc zabieg polegający na lokalnym powiększeniu kontrastu w bezpośrednim otoczeniu krawędzi (rys. 38). Efekt ten można osiągnąć stosując dwuwymiarowy cyfrowy filtr górnoprzepustowy. Uwypuklenie wysokich częstotliwości przestrzennych powoduje wzrost dostrzegalności drobnych szczegółów obrazu poprzez silniejsze zróżnicowanie jasności w otoczeniu konturów.

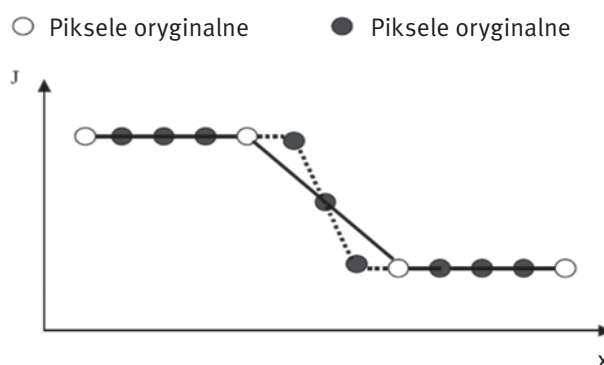
Innym sposobem poprawy ostrości jest zwiększenie stromości zboczy sygnału wizyjnego (rys.39). Stosując technikę nadpróbkowywania można utworzyć zbiór nowych pikseli w taki sposób, aby zrekonstruowany sygnał charakteryzował się pasmem charakterystycznym dla telewizji HDTV.

Do obliczenia wartości nowych pikseli są stosowane metody **interpolacji**. Proces interpolacji ma na celu utworzenie nowego, wcześniej nieistniejącego pikseli na podstawie pikseli sąsiadujących z pikselem tworzoną tak, aby był on jak najlepiej dopasowany optycznie do przetwarzanego obrazu. Dobierając właściwy algorytm interpolacji można osiągnąć efekt poprawy stromości zboczy bez zwiększania lokalnego kontrastu.





Rysunek 38.
Lokalne uwypuklenie konturów obrazu



Rysunek 39.
Technika nadpróbkowywania

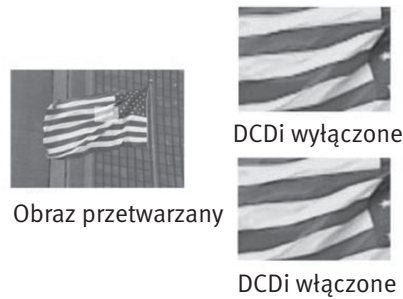
Przejścia pomiędzy fragmentami odpowiadającymi różnej jasności będą wtedy bardziej strome. W stosunku do wcześniej przedstawionej metody zwiększania ostrości, ta technika nie wprowadza zniekształceń grzbietu sygnału przed i po zboczu. Wspomniany sposób poprawy ostrości zastosowano w technologii D.I.S.T, stosowanej w niektórych odbiornikach HDTV, do zwiększenia rozdzielczości obrazu wizyjnego.

3.2.3 ALGORYTMY POPRAWY JAKOŚCI OBRAZU

Technologia D.I.S.T (ang. *Digital Image Scaling Technology*) opracowana została przez firmę JVC. Umożliwia ona redukcję migotania przy jednoczesnej poprawie rozdzielczości obrazu. Obraz przekazywany w konwencjonalnym 625-liniowym standardzie PAL z przeplotem zostaje na wstępie przetworzony do trybu progresywnego (czyli obraz jest wyświetlany bez przeplotu). Odbywa się to na drodze trójwymiarowej interpolacji wartości pikseli z linii półobrazów parzystego i nieparzystego, z wykorzystaniem relacji czasowych i przestrzennych między nimi. Specjalny algorytm interpolacji umożliwia uzyskanie wysokiej rozdzielczości w kierunku pionowym i podwojenie liczby linii w ramce do 1250. Sygnał wizyjny jest następnie formowany poprzez ekstrakcję 3 pól o częstotliwości 75 Hz z dwóch ramek 50 Hz i podawany na wyjście układu D.I.S.T. w trybie wybierania międzyliniowego 1250/75 Hz. Zwiększenie częstotliwości wyświetlania półobrazów, przyczynia się w tym przypadku do ograniczenia efektu migotania.

Redukcję zniekształceń krawędzi i linii umożliwia **technologia DCDi** (ang. *Directional Correlation Deinterlacing*) firmy Faroudja. W konwencjonalnej telewizji nieraz dostrzegalne są zniekształcenia polegające na poszarpaniu ukośnych linii lub konturów. Wrażenie to jest spotęgowane przy sekwencjach odtwarzanych w zwolnionym tempie (np. powtórka finiszu biegu – linie bieżni). Jednym z układów redukującym tego typu zniekształcenia jest DCDi. Ta technologia jest również wykorzystywana przez nadawców w USA do konwersji standardu NTSC do telewizji wysokiej rozdzielczości HDTV. Algorytm zaimplementowany w DCDi polega na inteligentnej interpolacji pikseli w zależności od charakteru ruchu obiektu w analizowanej scenie i kąta nachylenia konturów. Mechanizm interpolacji przebiega dzięki temu wzdłuż krawędzi nie dopuszczając do efektu ich poszarpania lub schodkowania, przy jednoczesnym zachowaniu ostrości i wierności oddania barw w miejscu przejść między kolorami.





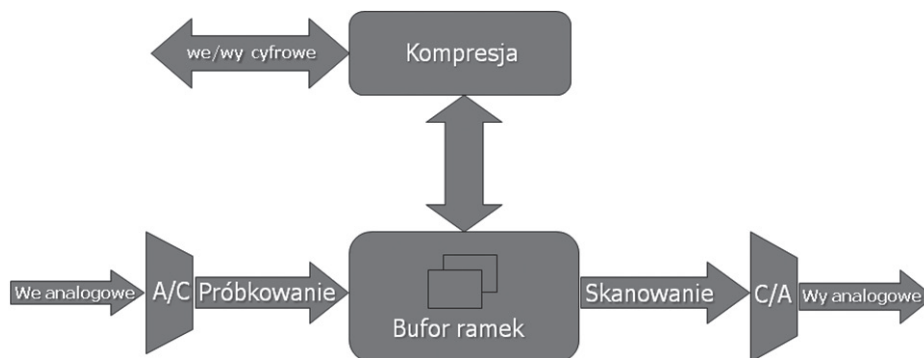
Rysunek 40.
Ilustracja działania DCDi

Na rys. 40 przedstawiono zastosowanie technologii DCDi do poprawy jakości obrazu. Z lewej strony znajduje się obraz flagi łopocząca na wietrze nadawany przez zwykłą telewizję. Jest to trudny obraz to wiernego odtworzenia w konwencjonalnej telewizji. Po prawej stronie przedstawiono powiększenia tego obrazu. W przypadku górnego obrazu filtry DCDi są wyłączone. Widać wyraźne poszarpanie krawędzi linii. W przypadku dolnego zaś DCDi jest włączona. Poszarpanie linii zniknęło, także połączenie obszarów czerwonych i białych jest bardziej naturalne.

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. Operacje przetwarzania sygnału wizyjnego w przykładowym układzie pozyskiwania obrazu można pogrupować w następujący sposób (rys. 41):

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.

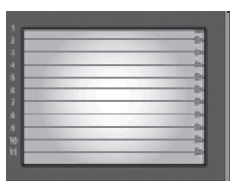


Rysunek 41.
Proces cyfrowego przetwarzania obrazów

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. 42) lub międzyliniowym (z przeplotem) (rys. 43). Stosowanie przeplotu jest charakterystyczne dla techniki telewizyjnej. Coraz częściej we współczesnych zastosowaniach wideo stosuje się jednak skanowanie progresywne, czyli tryb bez przeplotu.

Opisany wyżej proces cyfrowego przetwarzania obrazu jest realizowany np. w kamerze cyfrowej. Podczas kręcenia filmu kamerą, najpierw przetwornik (np. matryca CCD) próbkuje 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 42.

Odczyt zawartości bufora w trybie kolejnoliniowym (skanowanie progresywne)



Rysunek 43.

Odczyt zawartości bufora w trybie międzyliniowym (skanowanie międzyliniowe)

3.3 KOMPUTEROWA EDYCJA OBRAZU

Cyfrowe przetwarzanie obrazów wideo za pomocą 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 wyprowadzić go poprzez kartę do zewnętrznego magnetowidu. Taki sposób obróbki sekwencji obrazu nazywany jest montażem nieliniowym w odróżnieniu od klasycznego montażu przy pomocy 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 zniekształcenia. Warto jednak zauważyć, że pasmo sygnału nieskompresowanego: $720 \times 576 \times 25$ daje ok. 20 MB/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 3,7 MB/sek. na obraz, do czego dochodzi $2 \times 16 \times 48000 = 1,5$ MB/sek. na dźwięk, dając nam łącznie ok. 5 MB/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 kasy na dysk. Warto dodać, że kamera cyfrowa już w trakcie filmowania, aby zmniejszyć strumień danych, kompresuje obraz w stopniu 5:1 (kompresja DV). Niezaprzeczalną 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.

3.3.1 AUTHORIZING VCD I DVD

DVD authoring to tworzenie interfejsów płyt DVD, w formatach miniDVD, 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 są wykonywane 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,
- dostosowanie napisów i masek menu do wymogów Specyfikacji DVD-Video,
- połączenie wszystkich elementów audio-video odpowiednimi linkami w celu zapewnienia współpracy z użytkownikiem,
- sformatowanie obrazu płyty,
- zapisanie tego obrazu na odpowiednim nośniku (DVD-R, DLT, DVD-R for authoring, przenośny dysk twardy) i przekazanie go do tłoczni.

Wyróżnia się dwie klasy oprogramowania do DVD authoringu:

- narzędzia wbudowane w edytory wideo (niekiedy także edytory grafiki), gdzie DVD authoring jest ostatnim etapem pracy użytkownika, po zebraniu materiałów źródłowych i sporządzeniu prezentacji.
- wyspecjalizowane programy do DVD authoringu, które posługują się opracowanymi uprzednio w innych programach materiałami, np. przygotowanym filmem. Najbardziej popularny, profesjonalny program do DVD authoringu to Scenarist firmy Sonic.



Rozwój oprogramowania zamazuje częściowo różnice między klasami oprogramowania, gdyż programy do DVD authoringu dysponują zazwyczaj narzędziami do pobierania i prostej edycji materiałów filmowych, fotograficznych i dźwiękowych.

Poza tym określenie DVD authoring jest nieco mylące, gdyż mianem tym określa się technikę tworzenia interfejsu nie tylko płyty DVD, ale i VideoCD czy Super VideoCD, a także DVD on CD. Ta ostatnia, zwana miniDVD, umożliwia wykorzystanie na płycie CD niektórych technicznych walorów DVD.

Pierwszym stosowanym na szeroką skalę standardem zapisu materiału wizyjnego na popularnych nośnikach, jakimi są płyty CD, był standard VCD (ang. *Video Compact Disc*). Jest on standardem zapisywania na zwykłych płytach CD filmów i towarzyszących im ścieżek dźwiękowych. Standard ten, dzięki stosowanej kompresji MPEG, umożliwia zapisanie do 80 minut pełnoekranowego obrazu wideo wraz z towarzyszącym mu dobrej jakości dźwiękiem stereo na płycie o pojemności 700 MB. Aby pomieścić tak dużo materiału wizyjnego na pojedynczej płycie, w standardzie VCD jest stosowana kompresja zgodna ze standardem MPEG-1. W standardzie dodatkowo, oprócz zwykłej możliwości nagrywania materiału na płytę, istnieje również możliwość tworzenia i stosowania prostych bądź złożonych elementów menu, generacji rozdziałów, selekcji scen i innych elementów. Standard VCD umożliwia również przygotowywanie płyt będących prezentacją sekwencji zdjęć z możliwością dodania do nich podkładu dźwiękowego. Jakość przygotowanego materiału w standardzie VCD jest podobna do jakości materiału nagranych na kasecie VHS z tym, że obraz jest nieco bardziej rozmyty. Kolejnym krokiem w standardach zapisu materiału wizyjnego, będący jednocześnie następnikiem standardu VCD, jest standard SVCD (ang. *Super Video Compact Disc*). Standard ten umożliwia zapisanie na pojedynczej płycie CD o pojemności 700 MB do 60 minut bardzo dobrej jakości materiału wizyjnego z maksymalnie dwiema ścieżkami dźwiękowymi i czterema ścieżkami z napisami do filmu. Standard SVCD stosuje już kompresję w formacie MPEG-2. W standardzie również istnieje możliwość tworzenia i stosowania prostych bądź złożonych elementów menu, generacji rozdziałów oraz selekcji scen. Jakość materiału przygotowanego w standardzie SVCD jest znacznie lepsza niż w standardzie VCD, dzieje się tak z powodu wyższej rozdzielczości, w jakiej przygotowywany jest materiał na płytę. Jednak ostateczna jakość zależy od tego, jak dużo materiału musimy pomieścić na płycie – im więcej minut, tym mniejsza przepływność bitowa, a co za tym idzie gorsza jakość. Ostatnim z prezentowanych standardów zapisu materiału wizyjnego jest DVD (ang. *Digital Versatile/Video Disc*). Standard ten, dzięki zastosowanej standardu kompresji obrazu MPEG-2 umożliwia zapisanie na jednej płycie DVD o pojemności 4,7 GB nawet do 120 minut filmu o bardzo dobrej jakości, zawierającej również kilka ścieżek dźwiękowych (zarówno stereofonicznych jak i w standardzie dźwięku dookólnego Dolby Digital bądź DTS), a także zaawansowane systemy menu, napisy, zdjęcia oraz inne różnorodne elementy. Jeżeli zadowolimy się gorszą jakością, wykorzystując mniejsze przepływności bitowe i mniejsze rozdzielczości możemy zapisać nawet kilka godzin materiału z jakością zbliżoną do VHS, SVHS, VCD lub SVCD.

Format DVD jako pierwotnie pomyślany do zapisu pełnometrażowych filmowych z jakością odpowiadającą wymaganiom współczesnego zestawu kina domowego, musi również mieć różne dodatkowe właściwości, takie jak stosowanie wielojęzycznych ścieżek dialogowych (8 ścieżek dźwiękowych i 32 napisy dialogowe), stosowanie zabezpieczeń przed dziećmi, umożliwienie obserwacji materiału z różnych pozycji kamery (różne kąty widzenia) oraz dodatkowe funkcje interakcyjne.

Przygotowanie materiału wizyjnego do nagrania na płycie DVD bądź CD składa się z kilku etapów. Pierwszym etapem 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. Wszystko to umożliwi dokładne zaplanowanie przydziału bitów dla poszczególnych elementów materiału. Kolejnym etapem jest odpowiedni montaż oraz kompresja wszystkich materiałów, które chcemy wykorzystać do nagrania płyty. Mowa tu przede wszystkim o materiale wizyjnym, ale również na przykład o przygotowaniu ścieżek dźwiękowych. Etap ten może być dosyć złożony zważywszy, że zazwyczaj muszą być przygotowane ścieżki dźwiękowe w kilku wersjach językowych, oraz zrealizowane w różnych systemach. Standardowo, stosowany jest wielokanałowy dźwięk Dolby Digital, dwie lub trzy wersje Dolby Digital stereo z dubbingiem, a także jedna wersja dźwięku stereo PCM. Wszystkie wersje językowe muszą mieć jednakowe poziomy oraz barwę tak, by nie było słycać



dużych różnic w czasie przełączania pomiędzy różnymi wersjami językowymi. Ostatnim etapem produkcji 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. Na wstępie etapu authoringu należy dokładnie zaplanować elementy, które faktycznie chcemy umieścić w materiale. Efektem procesu authoringu jest przygotowanie oraz utworzenie odpowiedniej struktury formatu DVD, VCD bądź SVCD. Następnie w przypadku formatu DVD pliki na twardym dysku, zgodne z formatem DVD 1.0 formowane są zgodnie z formatem UDF–Bridge (*Universal Disc Format*), po czym może nastąpić przekopiowanie materiału na taśmę DLT (*Digital Linear Tape*), która jest wysyłana do producenta płyt DVD. Możliwe jest także bezpośrednie wypalenie obrazu płyty na nośnikach DVD-R lub DVD-RAM oraz ewentualnie CD-R.

Zadanie 3. Edycja sekwencji wizyjnej oraz eksport materiału wideo do formatu MPEG-2

Sprzęt potrzebny do wykonania zadania:

- komputer,
- oprogramowanie do montażu nieliniowego Pinnacle Studio 9, Adobe Premiere Elements 7.

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 są wyposażone w ścieżki wideo i audio, oś lub linię czasu oraz podstawowe narzędzia edycyjne. Edytowane ujęcia są umieszczane 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 tym trybie jest pokazywana 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 oferuje 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 dwumiarowymi są przenikania obrazu (miksy) oraz tzw. kurtyнки, natomiast efekty trójwymiarowe najczęściej są kojarzone 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 są przeznaczone 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, wyostżenia, 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 służy uzyskaniu profesjonalnego wyglądu tworzonego filmu.

b) Nakładanie scen

Kolejnym efektem jest nakładanie obrazów. 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 wykonać montaż wideo z zastosowaniem kroków przedstawionych w punkcie a) i b) niniejszej instrukcji.



c) Dodawanie ścieżki dźwiękowej

Produkcja materiału wizyjnego nie obędzie się bez dodatkowej ścieżki audio. W pakietach Pinnacle Studio 9 i Adobe Premiere Elements 7 jest możliwe zarówno nagrywanie własnej narracji, jak i zgrzywanie muzyki z płyt CD do pliku typu wav. Istnieje również możliwość dodawania zewnętrznych plików dźwiękowych.

Podobnie jak obraz wideo, również dźwięk może być poddawany różnym zabiegom edycyjnym. Podstawową operacją wykonywaną na ścieżkach audio jest ustawianieżądanego poziomu dźwięku. Ustawień dokonuje się za pomocą krzywych z obwiedniami opisującymi żądany poziom dźwięku. Warto tu zauważyć, że przy użyciu wspomnianych wcześniej efektów przejść również jest konieczna regulacja głośności zachodzących na siebie materiałów audio, co jest niezbędne do płynnego przejścia dźwięku z jednego materiału do drugiego. Do niedawna ustawiało się to ręcznie, natomiast obecnie większość dostępnych na rynku programów ma do tego specjalne narzędzia (ang. *crossfade*) lub też tworzy je automatycznie. Dodatkowo, w zaawansowanych programach jest dostępna również regulacja panoramy dźwięku, umożliwiająca precyzyjne balansowanie i ustawianie dźwięku w kanałach.

Niektóre programy miksują dźwięk z kilku (najczęściej 4 lub 8) ścieżek audio w czasie rzeczywistym. Należy jednak pamiętać, że znacznie to obciąża nie tylko procesor, ale co bardziej istotne, twarde dyski, które w trakcie miksowania muszą zebrać materiał znajdujący się w różnych obszarach. Dodatkowe efekty dostępne są dzięki specjalnym filtrom dźwięku, które możemy podzielić na dwie grupy: do korekcji dźwięku (korekcja pasma, barwy itp.) oraz do tworzenia efektów specjalnych (pogłosu, echa itp.). Wiele programów ma rozbudowane korektory (ang. *equalizers*) do precyzyjnego ustawienia żądanej barwy dźwięku.

Na etapie dodawania dodatkowej ścieżki dźwiękowej do wideo (punkt c) słuchacze powinni wykonać następujące operacje:

- nagrywanie własnej narracji,
- zgrzywanie muzyki z płyty CD,
- dodawanie zewnętrznych plików dźwiękowych,
- przycinanie plików dźwiękowych,
- dodawanie efektów wyciszania dźwięków.

Na końcu można również sprawdzić jakie możliwości edycji i montażu materiału wideo umożliwia program **Windows Movie Maker**.

4 ODTWARZANIE TREŚCI MULTIMEDIALNYCH

4.1 KINO DOMOWE

Kino domowe to zestaw urządzeń audio i wideo przeznaczony do oglądania filmów. 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 wielu 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. Dla 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 taśmy filmowej 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 20 Hz – 20 kHz, dwa kanały dookólne (lewy i prawy) działają w zakresie 80 Hz – 20 kHz, kanał LFE reprodukuje częstotliwości z pasma 20 Hz – 80 Hz. 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 rozpropagował się w kinie domowym. Stosowany w DTS algorytm kompresji CAC (*Coherent Acoustic Coding*) daje przy 20-bitowej rozdzielczości i próbkowaniu 44,1 kHz 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 44.1 kHz ze słowem 16-bitowym, zaproponowano próbkowanie 96 kHz 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 dekoderni o częstotliwości taktowania 48 kHz. Sygnał w formacie 96/24 jest dostępny na wyjściach cyfrowych odtwarzaczy i może być przetwarzany przez wysokiej klasy zewnętrzne przetworniki C/A.

4.2 AKUSTYKA POMIESZCZENIA

Dla 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 wytlumienie 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 niewytlumionych 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).

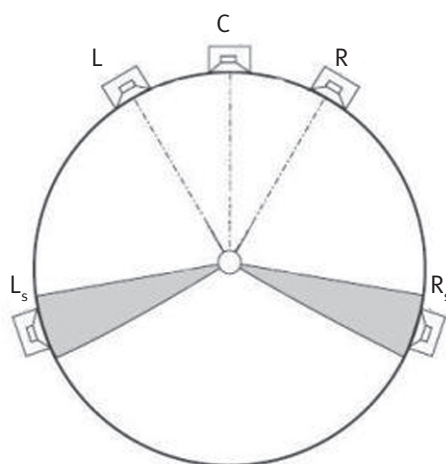
Jednym z najważniejszych pojęć determinujących akustykę pomieszczenia jest rezonans. W pomieszczeniu zamkniętym fale dźwiękowe odbijają się od ścian. W przypadku, gdy fale nachodzą na siebie niektóre częstotliwości mogą zostać wzmocnione a inne stłumione. Przy pewnych częstotliwościach nakładające się 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 na pierwszej pozycji 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, a 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 nas to





C – Kolumna Centralna
L,R – Kolumny przednie
L_s, R_s – Kolumny efektowe

Rysunek 44.

Rozmieszczenie głośników w standardzie 5.1

zupełnie nic, a w pozostałych przypadkach cena nie jest wysoka. Dlatego właśnie warto poświęcić czas, a czasami powierzchnię użytkową dla uzyskania optymalnych efektów.

Istnieje wiele metod ustawiania głośników. Sposób rozstawiania głośników w popularnym standardzie 5.1 zilustrowano na rys. 44.

Wszystkie trzy główne zestawy głośnikowe (L – lewy, C – centralny, R – prawy) powinny znajdować się w równej odległości od słuchacza. W przypadku, gdy fizycznie nie jest możliwe 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 posiada 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

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

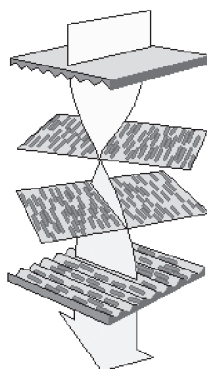
Głośniki średnio i wysoko tonowe powinny być widoczne przez słuchacza. Na drodze pomiędzy miejscem odsłuchowym a zestawami głośnikowymi nie powinny znajdować się przeszkody. Zasada ta nie dotyczy jednak głośnika niskotonowego (subwoofer). Wiąże się z nim zjawisko wszechkierunkowości. Oznacza to, że dźwięk 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 zasady 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.

4.3 WYŚWIETLACZE LCD

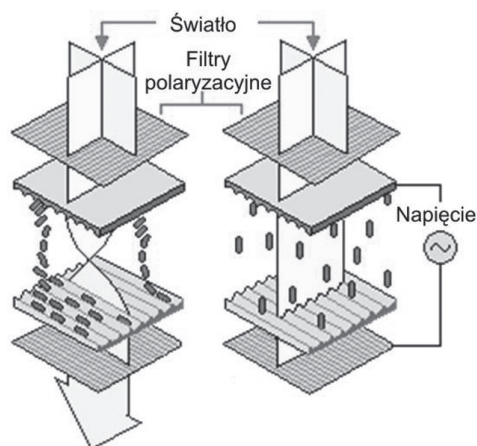
Ciekłe kryształy zostały wynalezione w XIX wieku przez austriackiego botanika Friedricha Reinitzera. Termin ciekły kryształ rozpropagował niemiecki fizyk Otto Lehmann. Ciekłe kryształy to substancje prawie przezroczyste, mające właściwości charakteryzujące zarówno ciała stałe, jak i ciecze. Światło przechodzące 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 45.
Zasada działania wyświetlacza LCD w technologii TN



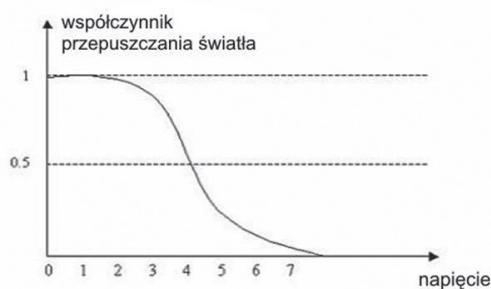


Rysunek 46.

Napięcie powoduje zmianę położenia molekuł ciekłego kryształu

W technologii TN ekran LCD składa się z dwóch warstw ciekłych kryształów umieszczonych pomiędzy dwiema odpowiednio wyprofilowanymi powierzchniami (rys. 45), 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. 46). Technologia TN jest najczęściej stosowana w niedrogich modelach monitorów komputerowych o niewielkich przekątnych obrazu – 15, 17 cali. Matryce tego typu nie najlepiej reprodukują 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.

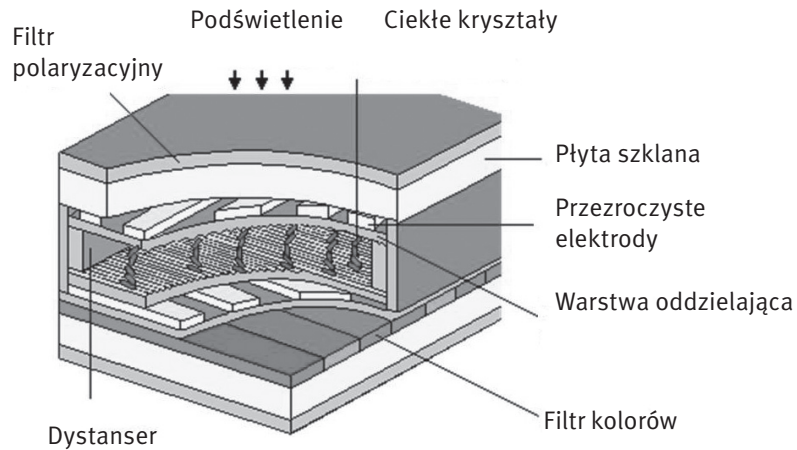
Jedną z wad paneli ciekłokrystalicznych jest gorsze odwzorowanie barw w stosunku do monitorów CRT. Ma to szczególnie znaczenie przy profesjonalnej obróbce zdjęć, zwłaszcza, gdy wyniki pracy mają trafić do drukarni. Normalny obserwator odróżnia od 300.000 do 1.000.000 barw natomiast przy pomocy techniki 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 przetwornika, aby uzyskać stany pośrednie (rys. 47). 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 47.

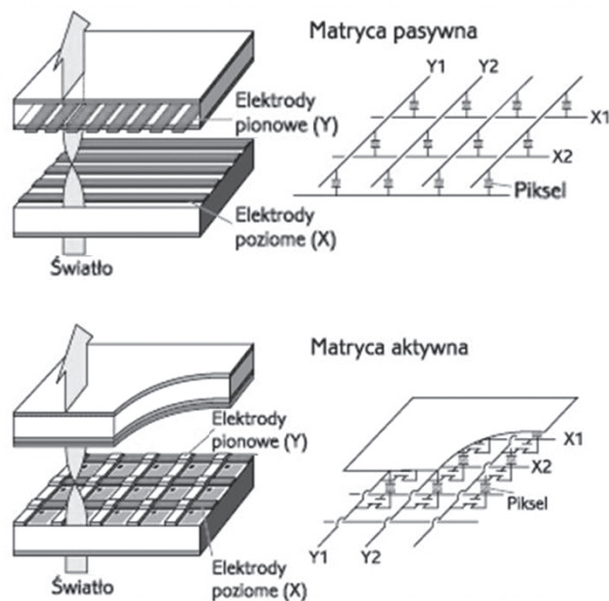
Różnicując napięcie na końcówkach ciekłego kryształu można modulować stopień zamknięcia przetwornika, aby uzyskać stany pośrednie

W **matrycach pasywnych** (rys. 48) ciekłe kryształy są adresowane poprzez ładunki lokalne, przy czym nic nie powstrzymuje ładunków elektrycznych przed rozptywaniem się na boki i wpływaniem na położenie molekuł kryształów sąsiednich. Adresowanie, czyli określenie pikselu, który ma być w danej chwili wystereowany, jest realizowane 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 48.
Budowa matrycy pasywnej

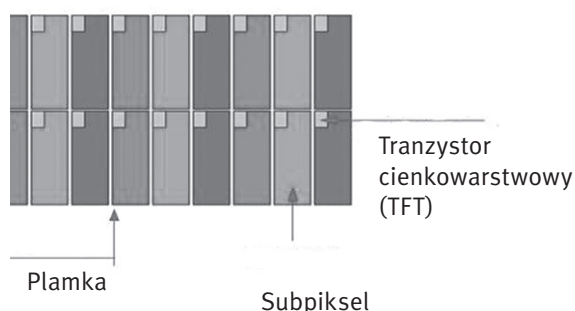
Matryce aktywne mają podobną budowę do matryc pasywnych. Podstawową różnicę stanowi warstwa tranzystorów cienkowarstwowych (ang. *thin film transistor*, TFT). Poprzez te tranzystory sterowane są kondensatory, które gromadzą i utrzymują ładunki elektryczne, zapobiegając ich rozptywaniu się na inne piksele (rys. 49).



Rysunek 49.
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 skręcone. 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.

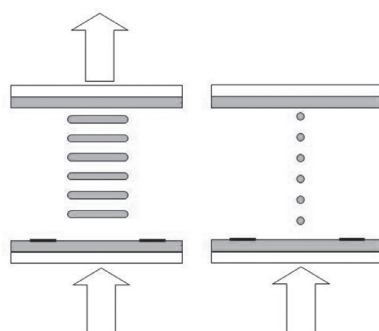
Specjalne filtry nadają kolor poszczególnym subpikselom ulokowanym na przedniej warstwie szkła (rys. 50). Trzy subpiksele, każdy w kolorze czerwonym, zielonym oraz niebieskim, formują piksel. Różne kombinacje kolorystyczne subpikseli dają obraz oraz kolor na ekranie.



Rysunek 50.
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 osiągnąć 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).



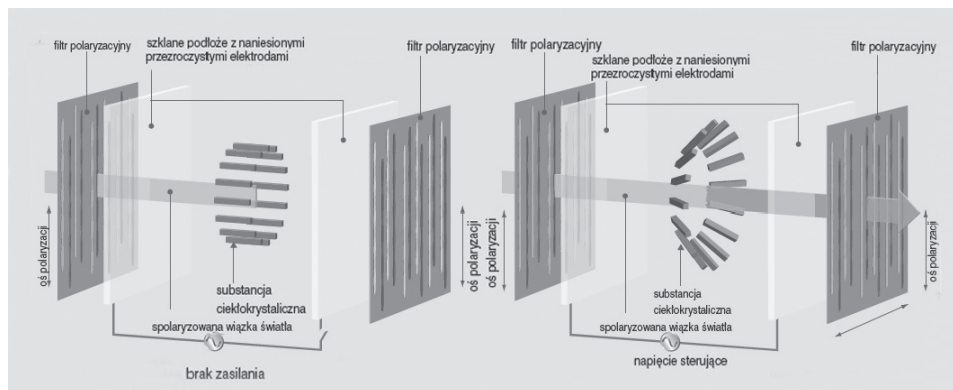
Rysunek 51.
Sposób ułożenia molekuł ciekłego kryształu w technologii IPS

Na rys. 51 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



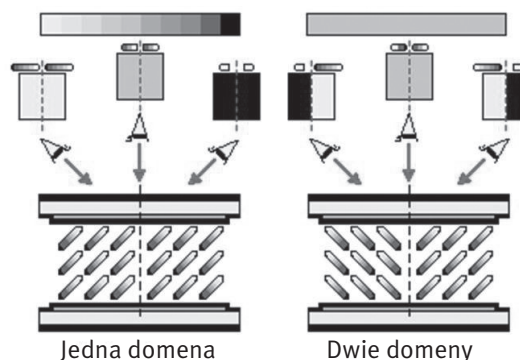
do siebie i do powierzchni wyświetlacza. Takie ułożenie molekuł ciekłego kryształu powoduje, że światło jest blokowane.

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. 52). Technologię PVA opracował 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 52.
Budowa wyświetlaczy LCD wykonanych w technologii MVA

Na rys. 53 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 molekuły ciekłego kryształu są ułożone prostopadłe do kierunku obserwacji, a zatrzymywane jest w przypadku równoległego ułożenia molekuł. W położeniach pośrednich przepuszczana jest tylko część światła. W przypadku jednej domeny (rys. 53 po lewej) obserwator przechodząc z lewej strony na prawą widzi stopniowy zanik świecenia pikseli. W przypadku zastosowania dwóch domen (rys. 53 po prawej) z lewej strony jedna domena przepuszcza światło, a druga je blokuje, z prawej strony natomiast 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.



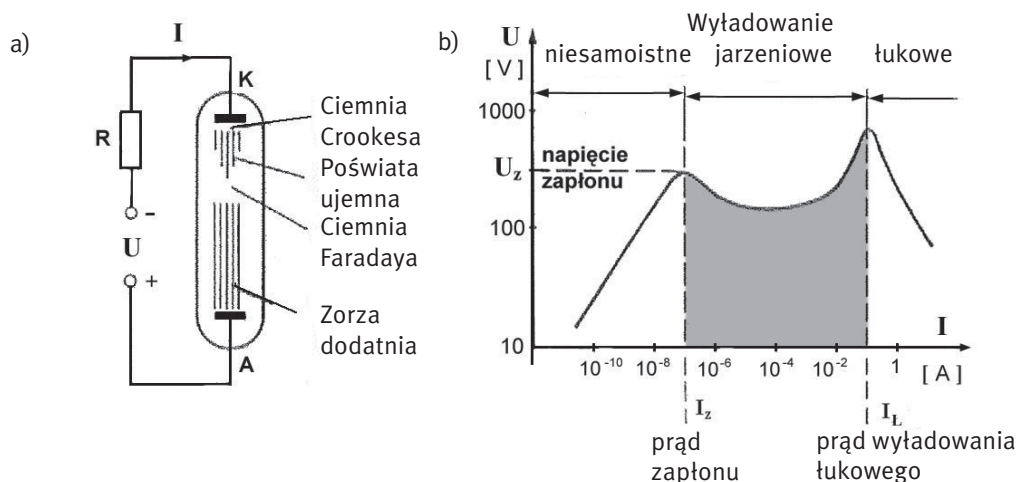
Rysunek 53.
Zasada działania matrycy MVA

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 60 Hz 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.

4.4 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 1 hPa) 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. 54) 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.

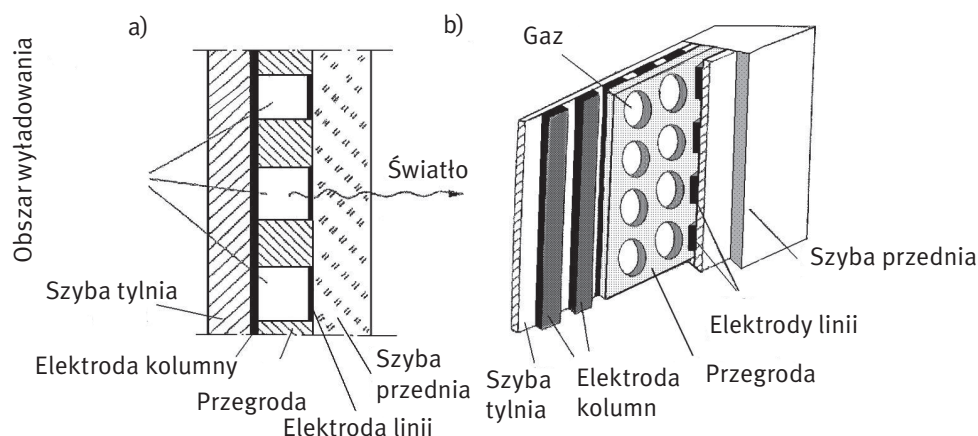


Rysunek 54.

Zjawisko wyładowania jarzeniowego w plazmie



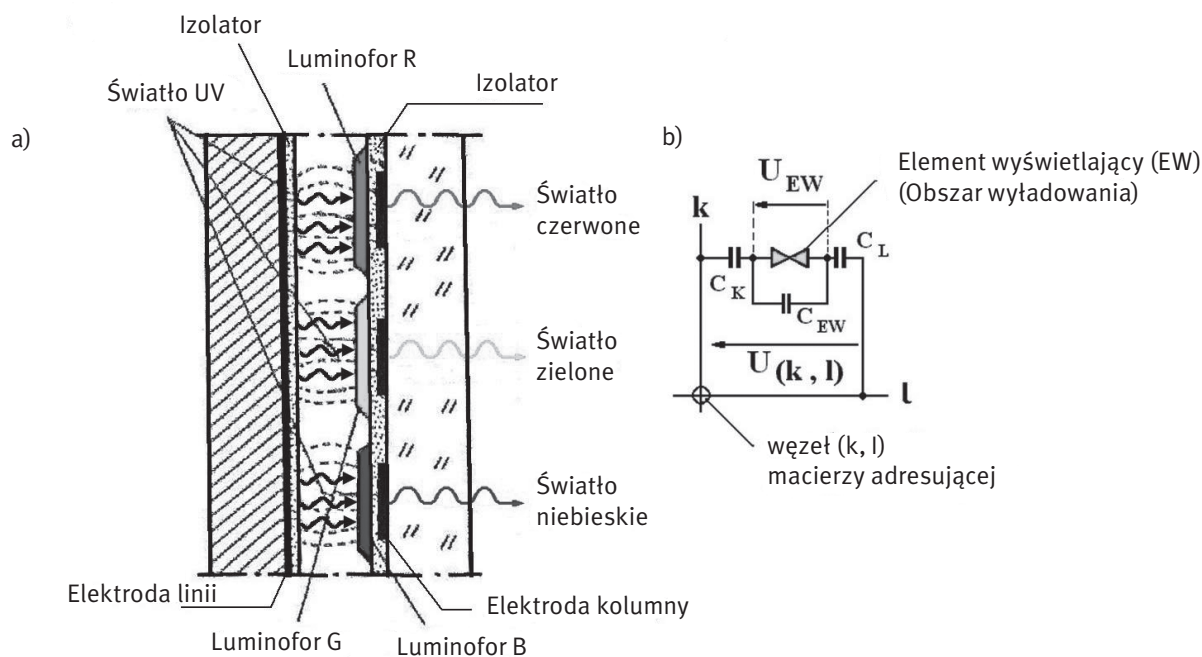
Na rys. 55 jest przedstawiona zasada konstrukcji najprostszego plazmowego wyświetlacza obrazów, wykorzystującego opisane wyżej zjawisko fizyczne. Dwa zestawy elektrod ułożone prostopadle względem siebie są naniesione na wewnętrzne powierzchnie szklanych płyt, tworzących obudowę przetwornika, pomiędzy którymi znajduje się rozrzedzony gaz. Adresowanie ekranu polega na dotknięciu 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 rozprzestrzenianiu 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 55.
Konstrukcja stałoprądowego wyświetlacza plazmowego

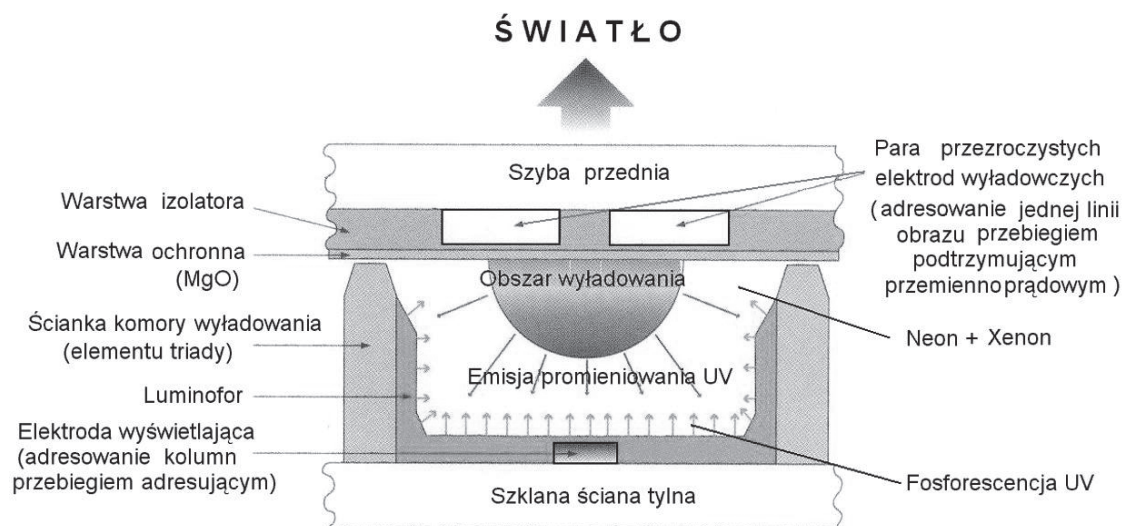
Wad stałoprądowych wyświetlaczy plazmowych nie ma **przemienno-prądowy wyświetlacz plazmowy**. Istota modyfikacji wobec ekranów DC PDP polega na odizolowaniu elektrod adresujących od gazu. Zapobiega to z jednej strony ich niszczeniu przez jony, z drugiej powoduje wtrącenie do obwodu wyładowania dwóch kondensatorów: C_K i C_L , tworzonych przez elektrody, izolator i obszar wyładowania (rys. 56).

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 zbyt dużym 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 przeladowania 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 dotarcia napięcia przemiennego do węzła



Rysunek 56.
Konstrukcja przemiennoprądowego wyświetlacza plazmowego

można generować pakiety błysków o zmiennej długości, sterując w ten sposób jasnością świecenia piksela. Od charakterystycznego sposobu pobudzenia obszaru wyładowczego do świecenia wyświetlacze stosujące opisaną wyżej zasadę wyświetlania noszą nazwę przemiennoprądowych ekranów plazmowych AC PDP. Do tej grupy rozwiązań zaliczają się wszystkie ekrany plazmowe, dostępne obecnie na rynku.



Rysunek 57.
Konstrukcja pojedynczego piksela wyświetlacza plazmowego

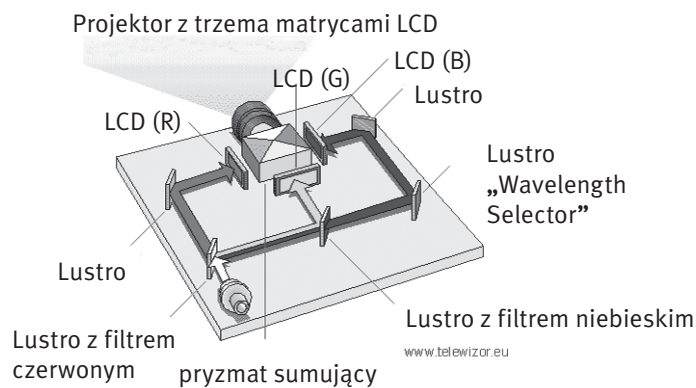
Na rys. 57 zilustrowano sposób uzyskiwania obrazów wielobarwnych przez ekrany plazmowe AC PDP. Wnętrze wyświetlacza wypełnione jest gazem lub mieszaniną gazów, świejących podczas wyładowania jarzeniowego światłem nadfioletowym (UV), które pobudza do świecenia paski luminoforów naniesione od wnętrza bańki. Umieszczenie luminoforów na ścieżce wyładowania powodowałoby ich niszczenie (wypalanie) przez jony świeącego gazu. Zjawisko to eliminuje stosowana w obecnie oferowanych rozwiązaniach wielobarw-



nych 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 świecącą plazmę.

4.5 PROJEKTORY MULTIMEDIALNE

Projektory 3LCD tworzone są 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. 58).



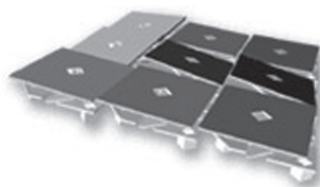
Rysunek 58.

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 luster (DLP chip), wynaleziony przez Dr. Larry Hornbeck z Texas Instruments w 1987 roku. W układ DLP wbudowana jest prostokątna macierz do 2 milionów mikroskopijnych luster. Zwierciadła są kwadratami o boku 16 mikronów, a przerwy między nimi nie mogą wynosić więcej niż 1 mikron (rys. 59).

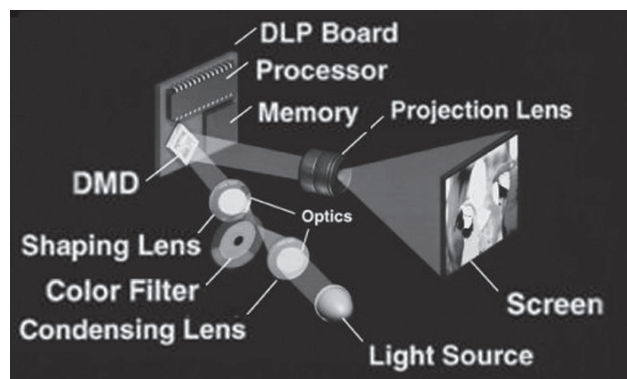


Rysunek 59.

Macierz luster

Mikrolustra układu DLP są umocowane 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 po-

wierzchni projekcyjnej. Strumień bitowy zakodowanego obrazu powoduje przełączanie 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. Obraz kolorowy w projektorach DLP jest tworzony w następujący sposób. Białe światło lampy projektora DLP przechodzi przez koło z filtrami kolorów trafia na mikrolustra układu DLP (rys. 60).



Rysunek 60.

Sposób tworzenia obrazu kolorowego w projektorze DLP z jednym Przetwornikiem

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 ma 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.

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 być 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. 61), wynikającego z chwilowego braku zbieżności barw.
- W najprostszyc konstrukcjach filtr składa się z trzech elementów, lecz coraz częściej nawet tanie projektory DLP posiadają 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 filtru mają mieć kształt nie prostokątny, lecz tezkowaty, nieco nachodzący na siebie.



Rysunek 61.
Efekt tęczy

Zadanie 4. 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 i Adobe Premiere Elements 7 stanowią kompletne zestawy 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.

Dodatkowo można wykorzystać wiele innych, dodatkowych możliwości oferowanych przez program, takich jak:

- a. załączanie różnorodnych dodatkowych danych w różnych formatach,
- b. import gotowych materiałów w formacie VOB lub MPG,
- c. dodawanie pokazu slajdów,
- d. automatyczne transkodowanie (najpierw tworzenie menu, potem konwersja).

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 pozwalają całkowicie przemodelować szablon, 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 authoring. 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 i Adobe Premiere Elements 7 większość tych operacji wykonywane jest automatycznie.

Końcowy etap to generowanie aplikacji – wypalanie płyty z gotową aplikacją. W przypadku VCD i SVCD można od razu nagrać płytę albo tylko utworzyć obraz płyty na dysku twardym do późniejszego wypalenia. W przypadku DVD można dodatkowo utworzyć foldery DVD na dysku, które mogą być uruchamiane przez odtwarzacze DVD, jak np. PowerDVD czy WinDVD. W zależności od karty oraz sprzętu wideo możliwy jest zapis w wielu formatach i standardach. Najczęściej stosowanymi są Video-CD, Super Video-CD oraz DVD-Video. W tym przypadku podstawą jest eksport materiału wideo wykorzystujący encodery formatu MPEG (dla Video-CD jest to MPEG-1, natomiast dla Super Video-CD i DVD-Video jest to MPEG-2). Wiele programów edycyjnych standardowo zostało wyposażonych w obsługę tych formatów, a w przypadku ich braku zawsze istnieje możliwość zainstalowania wtyczek.

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 i Adobe Premiere Elements 7 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.

Opracowanie wyników pracy

Efektom pierwszego etapu pracy powinien być gotowy zmontowany film ze wszystkimi niezbędnymi elementami. Do filmu należy dołączyć krótkie sprawozdanie, które powinno zawierać:

1. Krótką koncepcja filmu.
2. Spis wykorzystanych materiałów wizyjnych, dźwiękowych, graficznych.
3. Spis wykorzystanych efektów specjalnych.
4. Poszczególne wykonane kroki montażu materiału.

Efektom finalnym pracy powinna być przygotowana, a następnie wypalona płyta DVD z odpowiednim materiałem. W sprawozdaniu należy umieścić następujące informacje:

1. Sposób konwersji materiału wideo.
2. Sposób tworzenia menu (authoring).
3. Poszczególne kroki przygotowywania płyty 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 wyświetlanego na monitorze, telewizorze LCD i zaprezentowanego 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 2005



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

