

**Zaawansowane systemy baz danych - ZSBD**

## **Multimedialne bazy danych**

**Wykład prowadzi:  
Marek Wojciechowski**

**Multimedialne bazy danych**

Wykład poświęcony jest problematyce multimedialnych baz danych, ich specyfice oraz technikom wyszukiwania, składowania i prezentacji danych wykorzystywanym w ich kontekście.

Podstawowym celem niniejszego wykładu jest zapoznanie studenta z problematyką multimedialnych baz danych, ze szczególnym naciskiem na opis i wyszukiwanie treści multimedialnych.

Do zrozumienia treści wykładu niezbędna jest znajomość systemów baz danych i języka SQL.



## Plan wykładu

- Wprowadzenie: multimedia, podstawowe definicje
- Specyfika systemów multimedialnych baz danych
- Metadane
- Modele danych dla multimedialnych baz danych
- MPEG-7
- Przetwarzanie zapytań w multimedialnych bazach danych
- Wyszukiwanie w oparciu o zawartość
- Składowanie danych multimedialnych
- Prezentacja i transmisja danych multimedialnych

Multimedialne bazy danych (2)

Wykład rozpocznie się od wprowadzenia do multimedii i przedstawienia podstawowych definicji. Następnie wskazane będą cechy i wymagania stawiane systemom multimedialnych baz danych, które odróżniają je od tradycyjnych systemów baz danych. W kolejnej części wykładu omówione będą metadane, ich klasyfikacja i rola w opisie zawartości multimedialnej. Kolejny temat to modele danych dla multimedialnych baz danych. Następnie omówiony będzie standard opisu zawartości multimedialnej MPEG-7. Po nim przedstawiona będzie problematyka przetwarzania zapytań w multimedialnych bazach danych, ze szczególnym uwzględnieniem wyszukiwania obrazów w oparciu o zawartość (ang. Content-Based Image Retrieval). Na zakończenie, krótko nakreślone będą problemy składowania, prezentacji i transmisji danych multimedialnych, ze wskazaniem rozwiązań stosowanych w tym zakresie.



## Multimedia – podstawowe definicje

- **Obiekt medialny** – informacja zrozumiała dla komputera, zakodowana w jednym medium
  - media dyskretne: tekst, obraz
  - media ciągle: audio, wideo
- **Obiekt multimedialny** – informacja zrozumiała dla komputera, zakodowana w jednym lub więcej mediów
- **Dane multimedialne** – zbiór obiektów multimedialnych
- **System multimedialny** – system komputerowy wspierający wymianę informacji z użytkownikiem za pomocą kilku różnych mediów

Multimedialne bazy danych (3)

Obiekt medialny to informacja zrozumiała dla komputera, zakodowana w jednym medium. Wykorzystywane obecnie media to media wizualne i akustyczne: tekst, obrazy nieruchome, audio i wideo. Systemy wykorzystujące w komunikacji z człowiekiem zapach, smak i dotyk są na razie przedmiotem prac badawczych i prototypowych. Z punktu widzenia prezentacji multimedialnych istotny jest podział na media dyskretne (niezmiennie w czasie, statyczne), obejmujące tekst i obrazy nieruchome oraz media ciągle (zmiennie w czasie, dynamiczne), obejmujące audio i wideo.

Obiekt multimedialny to informacja zrozumiała dla komputera, zakodowana w jednym lub więcej mediów. Często dodawany jest jeszcze warunek, że co najmniej jedno medium powinno być nie-alfanumeryczne. W zasadzie termin „multimedia”, z racji przedrostka „multi-”, powinien być używany do opisu danych zakodowanych w więcej niż jednym medium. W praktyce przyjmuje się jednak, że obrazy czy muzyka w postaci cyfrowej również stanowią dane multimedialne, gdyż jest to zgodne z odczuciami użytkowników.

Dane multimedialne można najprościej zdefiniować jako zbiór obiektów multimedialnych.

System multimedialny (wg Meyera-Wegenera) to system komputerowy wspierający wymianę informacji z użytkownikiem za pomocą kilku różnych mediów. W zasadzie brak jest jednoznacznej, „ostrej”, powszechnie akceptowanej definicji systemu multimedialnego. Przykładowo, istnieje inna definicja mówiąca, że system multimedialny to system komputerowy umożliwiający w sposób zintegrowany tworzenie, przetwarzanie, prezentację i przesyłanie informacji zakodowanych w co najmniej jednym medium ciągłym i jednym dyskretnym.



## Charakterystyka typów multimedialnych (1/2)

- Obrazy
  - reprezentowane jako 2-wymiarowa tablica pikseli, gdzie wartość piksela reprezentuje jego kolor
  - zazwyczaj skompresowane (GIF, JPEG, PNG)
- Video
  - sekwencja klatek prezentowanych w odpowiednim tempie
  - podzielone na segmenty (fizyczne, logiczne)
  - zazwyczaj skompresowane (MPEG)
  - zazwyczaj uzupełnione o dane audio

Multimedialne bazy danych (4)

Podstawowe typy multimedialnych to obrazy nieruchome, dane wideo, dane audio i złożone obiekty multimedialne.

Obrazy są reprezentowane w postaci 2-wymiarowej tablicy pikseli (ang. pixel – picture element). Wartość przypisana danemu pikselowi reprezentuje jego kolor. Przykładowe liczby bitów przeznaczonych na zapamiętanie koloru to 8 dla obrazów w skali odcieni szarości i 24 dla obrazów kolorowych. Zazwyczaj obrazy są zapisane w postaci skompresowanej w celu redukcji ich rozmiaru. Przykładowe formaty to: GIF i PNG (kompresja bez utraty jakości) i JPEG (kompresja z utratą jakości, wykorzystująca właściwości ludzkiego zmysłu wzroku).

Dane wideo mają postać sekwencji klatek (ang. frame), gdzie każda klatka jest nieruchomym obrazem. Klatki muszą być prezentowane w odpowiednim ustalonym tempie, np. 25 klatek na sekundę. Sekwencja wideo może być podzielona na segmenty. Podstawowy podział to podział na segmenty fizyczne, gdzie każdy segment to sekwencja klatek pochodząca z jednego ujęcia kamery. Innym możliwym podziałem jest podział na segmenty logiczne reprezentujące fragmenty wideo o konkretnym znaczeniu. Dane wideo w postaci nieskompresowanej, ze względu na swój rozmiar, mogą być wykorzystywane jedynie na etapie profesjonalnej obróbki. Udostępniane użytkownikom w postaci cyfrowej są zawsze skompresowane. Kompresja wideo wykorzystuje przestrzenną i czasową nadmiarowość danych. Redukcja nadmiarowości przestrzennej polega na kompresji klatek metodą dla obrazów statycznych (np. JPEG). Redukcja nadmiarowości czasowej opiera się na reprezentacji niektórych klatek poprzez zapamiętanie zmian w stosunku do klatki poprzedzającej lub interpolację klatki z wykorzystaniem jednej klatki poprzedzającej i jednej następującej.

Podstawowe standardy kompresji wideo zostały opracowane przez grupę MPEG (Moving Pictures Experts Group): MPEG-1 zapewniający jakość porównywalną z VHS, wykorzystywany w VCD; MPEG-2 wykorzystywany w DVD i telewizji wysokiej rozdzielczości HDTV; MPEG-4 wykorzystywany w dystrybucji wideo poprzez Internet.

Dane wideo zazwyczaj uzupełnione są o dane audio, tworząc złożony obiekt multimedialny.



## Charakterystyka typów multimedialnych (2/2)

- Audio
  - cyfrowy zapis muzyki lub mowy
  - wynik próbkowania sygnału analogowego (PCM, MP3) lub dźwięk syntezowany (MIDI)
- Złożone dane multimedialne
  - filmy wideo z dźwiękiem (i opcjonalnie napisami)
  - prezentacje multimedialne
  - poszczególne media zapisane niezależnie lub razem tworząc nowy format

Multimedialne bazy danych (5)

Dane audio to muzyka i mowa. Mowa jest szczególnym przypadkiem danych audio, którym towarzyszą inne metadane niż muzyce i często również tekstowy zapis treści. Dane audio w postaci cyfrowej mogą być wynikiem próbkowania sygnału analogowego (np. z mikrofonu) albo być syntezowane na podstawie zapisu podobnego do nutowego. Dane będące wynikiem próbkowania mogą być zapisane w postaci nieskompresowanej (format PCM) lub skompresowanej (np. MP3 – kompresja z utratą jakości, wykorzystująca właściwości ludzkiego zmysłu słuchu). Najpopularniejszym formatem dla dźwięku syntezowanego jest MIDI.

Oprócz samodzielnych obrazów, audio i wideo, w praktyce często występują dane rzeczywiście multimedialne, czyli zakodowane w kilku różnych mediach. Typowym przykładem są filmy, w których sekwencji wideo towarzyszy zsynchronizowana z nią ścieżka dźwiękowa, a opcjonalnie też napisy (ang. subtitles). Innym przykładem danych integrujących wiele mediów są prezentacje multimedialne, gdzie tekstowo-obrazkowym slajdom towarzyszy dźwięk z komentarzem i opcjonalnie animacje. W przypadku złożonych obiektów multimedialnych, składające się na nie media mogą być składowane niezależnie lub razem, tworząc nowy format. Przykładem złożonego formatu jest choćby powszechnie znany format AVI, w którym przeplecione ze sobą są niezależnie kompresowane wideo i audio.



### Przyczyny gwałtownego rozwoju systemów multimedialnych

- Rzeczywistość postrzegana przez człowieka ma charakter multimedialny
- Dostępność mocy obliczeniowej
- Dostępność urządzeń do składowania danych o dużej pojemności
- Szybkie sieci komputerowe
- Efektywne standardy kodowania i transmisji
- Rozwój technologii mobilnych

Multimedialne bazy danych (6)

Podstawową motywacją dla systemów multimedialnych jest niepodważalny fakt, że rzeczywistość postrzegana przez człowieka ma charakter multimedialny. Na obraz rzeczywistości składają się informacje przekazane za pomocą wielu mediów, odbieranych za pomocą zmysłów (obraz, dźwięk, smak, zapach, dotyk). Zadaniem multimedialnych systemów informatycznych jest imitacja otaczającej człowieka rzeczywistości i sposobu jej postrzegania.

Jednakże dopiero od niedawna istnieją warunki do tworzenia systemów multimedialnych powszechnego użytku ze względu na:

- (a) dostępność tanich stacji roboczych zdolnych do przetwarzania danych multimedialnych (wydajne procesory i sprzętowe kodery/dekodery);
- (b) dostępność tanich urządzeń i nośników do składowania danych o dużej pojemności;
- (c) dostęp do szybkich sieci komputerowych (wzrost prędkości transmisji, rozwój infrastruktury, spadek cen nośników np. światłowodów, szybkie technologie transmisji danych np. ATM);
- (d) opracowanie efektywnych standardów kodowania i transmisji danych;
- (e) rozwój technologii mobilnych (telefony komórkowe umożliwiające odbiór zawartości multimedialnej, PDA, przewidywane nadejście UMTS).



## Zastosowania systemów multimedialnych

- Wideo na żądanie
- Muzyczne bazy danych
- Bazy danych z obrazami (nauka, sztuka, medycyna)
- Edukacja
- Walka z przestępczością i terroryzmem

Multimedialne bazy danych (7)

Przykładowe zastosowania systemów multimedialnych to:

(a) wideo na żądanie (ang. video on demand);

(b) muzyczne bazy danych;

(c) bazy danych z obrazami, przykładowo: kolekcje zdjęć rentgenowskich do wspierania diagnostyki medycznej, kolekcje fotografii dzieł malarskich, wirtualne galerie, itp.;

(d) edukacja: interaktywne kursy multimedialne, zdalne nauczanie, instrukcje obsługi, itp.;

(e) walka z przestępczością i terroryzmem: multimedialne bazy danych o przestępcach i terrorystach wykorzystane w połączeniu z infrastrukturą do monitoringu i technikami rozpoznawania obrazu, w tym ludzkich twarzy.

W większości zastosowań systemów multimedialnych mamy do czynienia z dużą kolekcją danych multimedialnych, które muszą być w sposób efektywny składowane, przeszukiwane i udostępniane. Stanowi to motywację dla ważnej klasy systemów multimedialnych – systemów multimedialnych baz danych.



## Systemy multimedialnych baz danych

$$\text{SMBD} = \text{SZMBD} + \text{MBD}$$

- SMBD – System multimedialnej bazy danych
- SZMBD – System zarządzania multimedialną bazą danych
  - wysoce wydajny SZBD
  - obsługujący dane alfanumeryczne i multimedialne
- MBD – Multimedialna baza danych
  - baza danych o dużej pojemności
  - zawierająca dane multimedialne
- SMBD a bazy danych z multimedialną zawartością

Multimedialne bazy danych (8)

System multimedialnej bazy danych (SMBD) składa się z bazy danych o dużej pojemności (MBD) z zawartością multimedialną i wysoce wydajnego systemu zarządzania bazą danych (SZBD), który wspiera i obsługuje obok alfanumerycznych typów danych również dane multimedialne w zakresie ich składowania, wyszukiwania i przetwarzania zapytań.

Należy zwrócić uwagę, że systemy oparte o bazę danych z multimedialną zawartością, nierealizujące funkcjonalności SZBD i wyszukiwania w oparciu o zawartość nie stanowią jeszcze systemów multimedialnych baz danych. Przykłady baz danych z multimedialną zawartością to:

- (a) płyty CD-ROM z multimediami,
- (b) katalogi obrazków z ikonkami prowadzącymi do obrazów w większej rozdzielczości,
- (c) systemy video on demand z wyszukiwaniem wg typowych słów kluczowych (tytuł, obsada, itp.),
- (d) systemy zarządzania dużymi zbiorami złożonych dokumentów.





### Systemy zarządzania multimedialną bazą danych (SZMBD)

- SZMBD =
  - SZBD
  - architektury do efektywnego składowania dużych wolumenów danych multimedialnych
  - techniki information retrieval (IR) – wyszukiwanie w oparciu o zawartość
- SZMBD integruje wiele technologii
- Podzbiorem funkcjonalności SZMBD jest funkcjonalność tradycyjnego SZBD

Multimedialne bazy danych (9)

System zarządzania multimedialną bazą danych (SZMBD) to wysoce wydajny system zarządzania bazą danych (SZBD), wykorzystujący architektury do efektywnego składowania dużych wolumenów danych multimedialnych oraz techniki Information Retrieval (IR) do wyszukiwania w oparciu o zawartość.

Systemy zarządzania multimedialnymi bazami danych integrują wiele technologii, w tym: bazy danych, Information Retrieval i przetwarzanie sygnałów. Podzbiorem ich funkcjonalności jest funkcjonalność tradycyjnych SZBD. W szczególności, system zarządzania multimedialną bazą danych, tak jak każdy system zarządzania bazą danych powinien oferować:

- (a) przetwarzanie transakcyjne,
- (b) zarządzanie współbieżnością,
- (c) odtwarzanie po awarii,
- (d) wersjonowanie danych,
- (e) ochronę zawartości bazy danych przed nieuprawnionymi użytkownikami i niedozwolonymi operacjami.



### Specyfika systemów multimedialnych baz danych

- Duży rozmiar danych
- Złożone struktury danych
- Zaawansowane przetwarzanie danych
- Zapytania różniące się od tradycyjnych
- Problemy prezentacji danych

Multimedialne bazy danych (10)

Właściwości, które sprawiają, że implementacja systemów zarządzania multimedialną bazą danych wymaga nowych rozwiązań znanych w systemach baz danych problemów to:

(a) duży rozmiar danych - Przykładowo, pojedynczy film wideo może zajmować kilka gigabajtów. Dlatego też systemy zarządzania multimedialną bazą danych oprócz składowania danych multimedialnych wewnątrz bazy danych w postaci dużych obiektów binarnych (ang. binary large object) muszą pozwalać na składowanie danych poza bazą danych (np. w systemie plików serwera, na serwerze WWW, na dedykowanych serwerach audio/wideo przygotowanych do transmisji strumieniowej). Duży rozmiar danych może też skłaniać do wykorzystania hierarchicznych struktur składowania oraz specjalnych organizacji danych na urządzeniach pamięci masowej.

(b) złożone struktury danych - Dane multimedialne są w praktyce uzupełnione opisem formatu i zawartości. Dla umożliwienia ścisłego wiązania metadanych z danymi multimedialnymi bardziej odpowiedni niż "czysty" model relacyjny jest obiektowy lub obiektowo-relacyjny model danych.

(c) zaawansowane przetwarzanie danych - Systemy zarządzania multimedialną bazą danych powinny realizować takie operacje jak np. konwersja formatu, skalowanie i inne transformacje obrazów. Dla wygody użytkowników i bezpieczeństwa danych ważne jest, aby realizacja tego typu operacji nie wymagała wydobycia danych z bazy danych i wykorzystania do ich obróbki innego oprogramowania.



### Specyfika systemów multimedialnych baz danych (cd.)

- Duży rozmiar danych
- Złożone struktury danych
- Zaawansowane przetwarzanie danych
- Zapytania różniące się od tradycyjnych
- Problemy prezentacji danych

Multimedialne bazy danych (11)

(d) zapytania różniące się od tradycyjnych - Charakterystyczne dla multimedialnych baz danych jest wyszukiwanie oparte o zawartość, rozumiane jako wyszukiwanie, którego kryteria nie odnoszą się do zapisu danych w bazie danych, ale do właściwości wizualnych lub akustycznych, które muszą być wyekstrahowane z obiektu multimedialnego. Kryteria wyszukiwania w oparciu o zawartość najczęściej są specyfikowane poprzez podanie przykładu. Charakterystyczne dla wyszukiwania obiektów podobnych są: możliwość przypisywania wag poszczególnym właściwościom oraz tolerowanie przybliżonych lub niekompletnych wyników.

(e) problemy prezentacji danych - Problem ten dotyczy danych wrażliwych na opóźnienia, czyli sekwencji audio i wideo. Występuje w związku z dużym rozmiarem danych, ograniczoną przepustowością serwera i kanałów transmisyjnych oraz ograniczonym buforem danych klienta. W szczególności dane wideo powinny być prezentowane w oparciu o transmisję strumieniową. Serwer musi być więc przygotowany do strumieniowej transmisji danych, ze współbieżną obsługą wielu klientów z zachowaniem odpowiedniej jakości usług.



### Systemy rozproszonych multimedialnych baz danych

- **Motywacje:**
  - wymagania pamięci i przepustowości sieci
  - często naturalne rozproszenie
- **Zalety systemów rozproszonych:**
  - efektywność systemu (moc obliczeniowa, większe pojemności, równoleglenie pracy)
  - krótsza ścieżka od serwera do użytkownika
- **Klasyczne problemy rozproszonych baz danych**

Multimedialne bazy danych (12)

W kontekście systemów multimedialnych (w tym multimedialnych baz danych) często rozważa się architektury rozproszone. Są dwie główne motywacje dla multimedialnych systemów rozproszonych. Po pierwsze, systemy scentralizowane wymagają olbrzymich pamięci masowych do składowania danych i dużych przepustowości sieci do ich przesyłania. Szczególnie dla baz danych wideo może się okazać, że scentralizowany system udostępniania filmów wideo przy aktualnym stanie technologii i dostępnej infrastruktury dla konkretnego planowanego obciążenia nie będzie mógł zapewnić odpowiedniej jakości usług. Po drugie, często dane multimedialne stanowiące multimedialną bazę danych są naturalnie rozproszone np. gdy poszczególne typy multimedii (audio, wideo, obrazy) są składowane na serwerach dedykowanych do udostępniania konkretnego typu danych.

Zaletą systemów rozproszonych to możliwość zwiększania efektywności i pojemności systemu poprzez dodawanie nowych węzłów. Pozwala to zrównoleglić pracę i rozłożyć obciążenie sieci. Ponadto, w połączeniu z odpowiednim rozmieszczeniem i replikacją danych na poszczególnych węzłach systemu można skrócić ścieżkę między serwerem a użytkownikiem.

Systemy rozproszonych multimedialnych baz danych wymagają jednak rozwiązywania problemów znanych z klasycznych systemów rozproszonych baz danych takich jak:

- (a) wybór sposobu rozproszenia (alokacji) danych między serwery,
- (b) wybór danych, które mają być replikowane,
- (c) integracja danych z różnych serwerów przy przetwarzaniu zapytań,
- (d) zarządzanie rozproszonymi transakcjami,
- (e) pielęgnacja rozproszonych metadanych.



## Metadane

- Metadane sygnałowe
  - wyodrębnione właściwości sygnału stanowiącego zapis zawartości multimedialnej
- Metadane semantyczne
  - informacje o obiektach, postaciach, zdarzeniach, itp. reprezentowanych przez treści multimedialne
- Metadane zewnętrzne
  - informacje nieodnoszące się bezpośrednio do zawartości medium, tylko do sposobu jego tworzenia, przechowywania, udostępniania, klasyfikacji

Multimedialne bazy danych (13)

Metadane to „dane o danych”. Metadane mają szczególne znaczenie w opisie i wyszukiwaniu danych multimedialnych. Zależnie od ich charakteru, niektóre metadane mogą być automatycznie wyekstrahowane z zawartości binarnej, a niektóre muszą być wprowadzone przez człowieka i towarzyszyć zawartości multimedialnej. Ze względu na charakter metadanych można pokusić się o ich następującą klasyfikację:

Metadane sygnałowe to wyodrębnione właściwości sygnału stanowiącego zapis zawartości multimedialnej. Mogą być stosunkowo łatwo wygenerowane automatycznie w oparciu o specjalne algorytmy. Przykłady metadanych sygnałowych to: histogram kolorów czy tekstura dla obrazów, dynamika ruchu dla sekwencji wideo, melodia dla sekwencji audio.

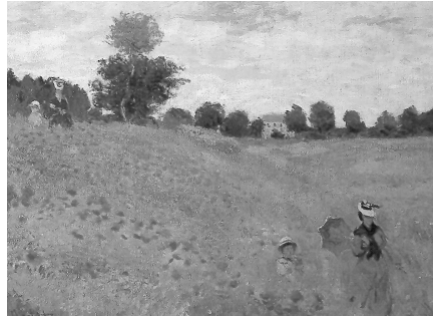
Metadane semantyczne to metadane opisujące znaczenie danych, czyli zawierające informacje o obiektach, postaciach, zdarzeniach, itp. reprezentowanych przez treści multimedialne.

Metadane semantyczne są trudne bądź niemożliwe do w pełni automatycznego uzyskania. Powstają one w wyniku tworzenia adnotacji przez człowieka, który może być wspomagany przez narzędzia dokonujące automatycznej segmentacji obrazu, podziału sekwencji wideo na ujęcia itp.

Metadane zewnętrzne to informacje nieodnoszące się bezpośrednio do zawartości medium, tylko do sposobu jego tworzenia, przechowywania, udostępniania, czy klasyfikacji. Przykłady metadanych zewnętrznych to reżyser filmu, nazwa pliku źródłowego, typ kompresji pliku, tytuł zdjęcia, czas i miejsce wykonania zdjęcia.



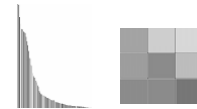
## Rodzaje metadanych - Przykład



### Metadane zewnętrzne:

Autor: Claude Monet  
Tytuł: Maki w pobliżu Argenteuil  
Styl: impresjonizm  
Format: jpeg  
Rozmiar: 640 x 480

### Metadane sygnałowe:



### Metadane semantyczne:

łąka pełna maków  
na dole z prawej kobieta z parasolką i dziecko

Multimedialne bazy danych (14)

Slajd przedstawia przykłady poszczególnych rodzajów metadanych dla fotografii dzieła malarskiego w postaci cyfrowej.

Metadane sygnałowe w tym wypadku to histogram reprezentujący udział kolorów w obrazie i macierz pokazująca dominujące kolory w poszczególnych regionach obrazu. Metadane semantyczne obejmują ogólny opis zawartości („łąka pełna maków”) oraz opis najważniejszych obiektów wraz z ich położeniem („na dole z prawej kobieta z parasolką i dziecko”). Metadane zewnętrzne to informacje o autorze, tytule i stylu dzieła oraz informacje o cyfrowej reprezentacji tj. format graficzny i rozmiar w pikselach.



### Modele danych dla multimedialnych baz danych

- Model relacyjny
  - nieodpowiedni
  - problemy z reprezentacją złożonych metadanych i wiązaniem ich z zawartością binarną
- Model obiektowy
  - odpowiedni, elastyczny, ale nie odniósł sukcesu
- Model obiektowo-relacyjny
  - odpowiedni, praktyczny kompromis
- Model semistrukturalny (XML)
  - odpowiedni do reprezentacji metadanych

Multimedialne bazy danych (15)

Modele danych wykorzystywane w dzisiejszych systemach baz danych to model relacyjny, model obiektowy, model obiektowo-relacyjny i model semistrukturalny.

Model relacyjny to model, w którym struktury danych mają postać zbioru wzajemnie ze sobą powiązanych relacji (tabel). Czysty model relacyjny nie przewidywał obsługi danych binarnych. Obecne systemy relacyjne obok alfanumerycznych typów danych, oferują również typ BLOB (Binary Large Object), umożliwiający składowanie dużych obiektów binarnych w komórce tabeli. Mimo tego, model relacyjny nie jest odpowiedni do reprezentacji danych multimedialnych, gdyż oferowana przez niego organizacja danych nie pozwala na reprezentację złożonych metadanych semantycznych. Nie jest możliwe również ściśle wiązanie metadanych z zawartością binarną.

Model obiektowy stanowi odpowiedź na ubóstwo modelu relacyjnego. Umożliwia modelowanie złożonych struktur i obiektów oraz wiąże dane z operacjami na nich. Model ten jest potencjalnie atrakcyjny dla multimedii: obiekt multimedialny może być reprezentowany poprzez obiekt w bazie danych, którego jednym z atrybutów jest zawartość binarna, a pozostałe reprezentują wszelkiego rodzaju metadane. Problemem modelu obiektowego jest to, że nie odniósł dużego sukcesu gdyż mimo 20 lat rozwoju ciągle nie zaproponowano satysfakcjonujących rozwiązań problemów: współbieżnego dostępu, optymalizacji zapytań, odtwarzania po awarii, itp.



### Modele danych dla multimedialnych baz danych (cd.)

- Model relacyjny
  - nieodpowiedni
  - problemy z reprezentacją złożonych metadanych i wiązaniem ich z zawartością binarną
- Model obiektowy
  - odpowiedni, elastyczny, ale nie odniósł sukcesu
- Model obiektowo-relacyjny
  - odpowiedni, praktyczny kompromis
- Model semistrukturalny (XML)
  - odpowiedni do reprezentacji metadanych

Multimedialne bazy danych (16)

Model obiektowo-relacyjny to odpowiedź „świata relacyjnego” na model obiektowy. Doczekał się standaryzacji wraz z wersją standardu SQL99. Jest to model relacyjny uzupełniony o: złożone typy danych, dziedziczenie, kolekcje, referencje do obiektów, itp. Model obiektowo-relacyjny jest powszechnie implementowany gdyż większość systemów relacyjnych przeszła ewolucję stając się systemami obiektowo-relacyjnymi. Model ten jest odpowiedni dla multimedialnych baz danych i stał się podstawą dla standardu SQL/MM, standaryzującego sposób reprezentacji i przetwarzania zaawansowanych rodzajów danych (w tym multimediiów) w bazach danych.

Model semistrukturalny umożliwia reprezentację danych posiadających pewną strukturę, ale niekoniecznie na tyle określoną, aby odpowiedni dla nich był model relacyjny lub obiektowy. Podstawowym formatem reprezentacji danych semistrukturalnych jest obecnie XML. W kontekście multimedialnych baz danych, XML jest atrakcyjnym formatem zapisu metadanych. Może być więc wykorzystany w multimedialnych bazach danych w połączeniu z modelem relacyjnym, obiektowym lub obiektowo-relacyjnym. Format XML stał się podstawą dla standardu reprezentacji metadanych o danych multimedialnych MPEG-7.





### Modele danych dla multimedialnych baz danych – Stan aktualny

- Systemy komercyjne (Oracle, IBM DB2):
  - wykorzystanie rozszerzeń obiektowo-relacyjnych
  - rozwiązania firmowe dla metadanych sygnałowych
  - prawie brak wsparcia dla metadanych semantycznych
- Standardy:
  - SQL/MM – model obiektowo-relacyjny; brak wsparcia dla metadanych semantycznych
  - MPEG-7 – reprezentacja wszystkich typów metadanych w postaci XML; na razie słabe wsparcie przez systemy zarządzania bazą danych

Multimedialne bazy danych (17)

Wiodące systemy zarządzania bazami danych ogólnego przeznaczenia (Oracle i IBM DB2) obsługują dane multimedialne w oparciu o mechanizmy obiektowo-relacyjne, a przede wszystkim mechanizm definiowania złożonych typów danych. Reprezentacja i ekstrakcja metadanych sygnałowych oraz wyszukiwanie w oparciu o zawartość są realizowane tylko dla obrazów, w oparciu o rozwiązania firmowe, przy czym Oracle jako alternatywę dla swoich rozwiązań oferuje interfejs zgodny ze standardem SQL/MM. Żaden z systemów zarządzania bazą danych ogólnego przeznaczenia nie wspiera w szczególności sposobu reprezentacji metadanych semantycznych. Ogólne wsparcie sprowadza się do możliwości definiowania złożonych struktur obiektowych i wsparcia dla XML. Metadane zewnętrzne nie stanowią problemu dla istniejących systemów zarządzania bazami danych, gdyż jako proste dane alfanumeryczne mogą być reprezentowane w strukturach relacyjnych i obiektowo-relacyjnych.

Obowiązujące standardy dla multimedialnych baz danych to SQL/MM i MPEG-7, rozwijane niezależnie przez dwa różne środowiska.

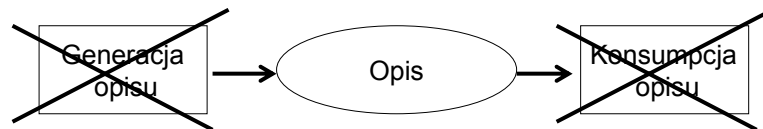
SQL/MM jest oparty o model obiektowo-relacyjny i uzupełnia SQL o typy danych do reprezentacji zaawansowanych rodzajów danych. Z multimediiów w tej chwili obejmuje jedynie obrazy, przewidując dla nich ograniczony zestaw metadanych zewnętrznych, podstawowe metadane sygnałowe i wyszukiwanie w oparciu o zawartość z ich uwzględnieniem. Standard SQL/MM nie porusza kwestii metadanych semantycznych.

MPEG-7 standaryzuje sposób reprezentacji wszelkiego rodzaju metadanych o wszelkiego typu danych multimedialnych i jest oparty o XML. Nie standaryzuje metod ekstrakcji i wykorzystywania metadanych. Na razie MPEG-7 jest słabo wspierany w systemach zarządzania bazą danych. Wsparcie dla opisów MPEG-7 nie wykracza w nich poza ogólne wsparcie dla XML.



## MPEG-7

- MPEG-7 = Multimedia Content Description Interface
- Standard normujący opisy zawartości multimedialnej



- Adresatami opisu mają być ludzie lub programy komputerowe
- Przewidywane zastosowania:
  - wyszukiwanie w multimedialnych bazach danych
  - filtracja treści multimedialnych

Multimedialne bazy danych (18)

W przeciwieństwie do standardów MPEG-1, MPEG-2 i MPEG-4 dotyczących reprezentacji (sposobu kodowania) zawartości obiektów multimedialnych, standard MPEG-7 (Multimedia Content Description Interface) dotyczy opisu zawartości multimedialnej. Standard MPEG-7 normuje jedynie samą formę opisu, nie standaryzując technik generacji i konsumpcji opisów. Przyczyny wyłączenia generacji opisów, ekstrakcji właściwości oraz metod wykorzystywania opisów z zakresu standardu były następujące:

- (a) mechanizmy te mają stanowić obszar swobodnej konkurencji, sprzyjającej rozwojowi technologicznemu;
- (b) standaryzacja algorytmów nie jest konieczna dla możliwości współdziałania różnych systemów;
- (c) charakter i funkcjonalność wyszukiwarek zwykle zależy od zastosowania, a format MPEG-7 ma z założenia być niezależny od zastosowań.

Zakłada się, że odbiorcami opisów MPEG-7 będą zarówno ludzie jak i programy komputerowe oraz inteligentne urządzenia elektroniczne. Opisy MPEG-7 powinny towarzyszyć danym multimedialnym składowanym w bazach danych, ale również udostępnianym przez radio, telewizję, media. Przewidywane zastosowania MPEG-7 to:

- (a) wyszukiwanie informacji w multimedialnych bazach danych;
- (b) filtracja treści multimedialnych odbieranych za pomocą telewizji, radia itp. w oparciu o opisy MPEG-7 przesyłane wraz z zawartością i ustawienia preferencji użytkownika.



## Założenia standardu MPEG-7

- Standaryzacja opisu metadanych
- Niezależność od sposobu kodowania medium
- Elastyczność
- Rozszerzalność
- Oparcie o język XML
  - język definicji składni (DDL) oparty o XML Schema
- Binarny format BiM do kompresji i transmisji opisów

Multimedialne bazy danych (19)

Podstawowym założeniem przyświecającym twórcom standardu MPEG-7 była standaryzacja opisu metadanych. Wcześniej pojawiało się wiele niezgodnych ze sobą propozycji formatów opisów multimedialnych, MPEG-7 korzysta z ich doświadczeń i normuje format opisu jako powszechnie uznany i oficjalny standard.

MPEG-7 jest całkowicie niezależny od sposobu kodowania i przechowywania medium i może dotyczyć np. obrazów narysowanych na papierze. Jeśli jednak dany format kodowania może zawierać informacje o strukturalnej dekompozycji obiektu multimedialnego (jak np. MPEG-4), informacje te mogą być wykorzystane przy tworzeniu opisów MPEG-7.

MPEG-7 jest standardem elastycznym i rozszerzalnym. Jako efekt analizy szerokiego spektrum potencjalnych zastosowań, MPEG-7 jest formatem ogólnego przeznaczenia. Jednocześnie daje on możliwość rozszerzenia składni języka metadanych o elementy przydatne w specyficznych zastosowaniach.

MPEG-7 jest oparty o język XML. Wybór formatu do tekstowych opisów był w zasadzie oczywisty, biorąc pod uwagę elastyczność, prostotę i pozycję na rynku języka XML. Dzięki temu, że twórcy MPEG-7 postanowili nie opracowywać od podstaw nowego formatu, ale oparli się o XML, automatycznie dostępnych jest szereg narzędzi do parsowania i przeszukiwania opisów. Język definicji składni opisów MPEG-7 (Description Definition Language - DDL) został zdefiniowany w oparciu o XML Schema, który jest silnym narzędziem opisu struktury (gramatyki) dokumentów XML.

Mimo niekwestionowanych zalet XML, ma on też pewne wady. XML jest nieodpowiedni dla transmisji strumieniowej, choćby ze względu na nieoszczędny sposób reprezentacji informacji. Dlatego też, w ramach standardu MPEG-7 z myślą o transmisji (szczególnie strumieniowej) opisów opracowany został binarny format opisów MPEG-7 o nazwie BiM.

BiM umożliwia bardzo dobrą kompresję opisów XML-owych z zachowaniem pewnych możliwości nawigacji po strukturze opisu na poziomie binarnym.



## Zawartość opisów MPEG-7

- Zbudowane z deskryptorów i schematów opisu
- Dekompozycja przestrzenna i czasowa
  - segmenty, regiony
- Opisy znaczenia zawartości (metadane semantyczne)
  - opisy zdarzeń, osób, obiektów, czynności
- Niskopoziomowe opisy zawartości (metadane sygnałowe)
  - kształt, tekstura, kolor, lokalizacja, trajektoria ruchu, melodia
- Informacje uzupełniające (metadane zewnętrzne)
  - forma, dostępność materiału, klasyfikacja, kontekst

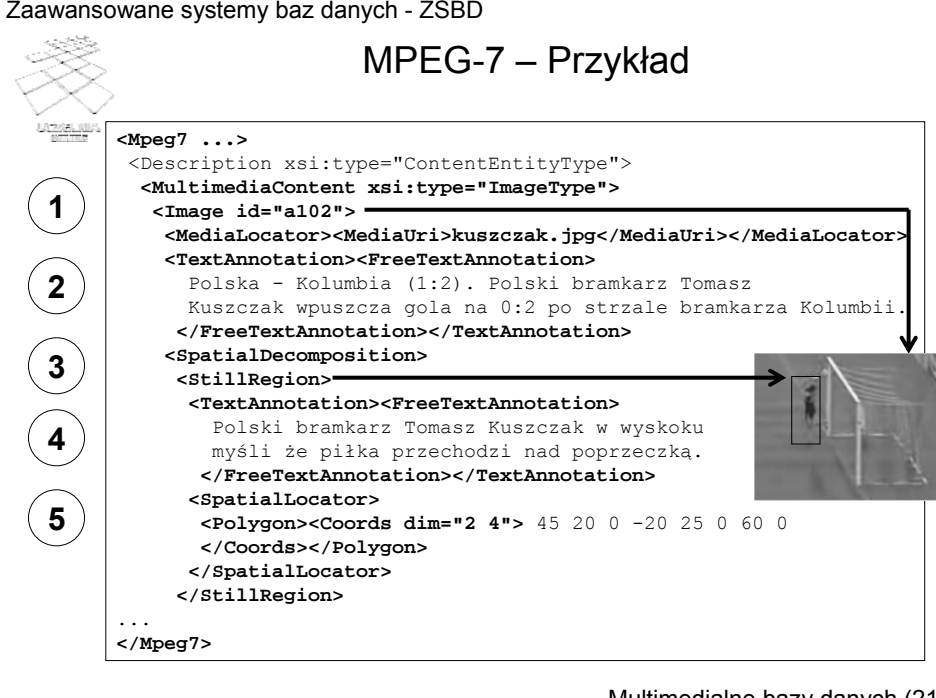
Multimedialne bazy danych (20)

Od strony technicznej opisy MPEG-7 są tworzone za pomocą deskryptorów (reprezentujących poszczególne właściwości obiektów multimedialnych) i schematów opisu (specyfikujących strukturę i semantykę związków między elementami opisu tj. deskryptorami i innymi schematami opisu).

Treść dokumentów MPEG-7 oprócz ogólnego opisu całego obiektu multimedialnego, zawiera typowo opisy poszczególnych elementów składowych obiektu, wyodrębnionych w drodze dekompozycji przestrzennej i/lub czasowej. Przykładem dekompozycji czasowej jest podział sekwencji wideo na segmenty. Z kolei dekompozycja przestrzenna może dotyczyć samodzielnych obrazów nieruchomych lub klatek filmu wideo i polega na wskazaniu na obrazie regionów zawierających interesujące obiekty.

Opisy MPEG-7 mogą obejmować wszystkie rodzaje metadanych. Za pomocą MPEG-7 można opisać semantykę obiektu i jego elementów (opisy zdarzeń, osób, obiektów, czynności) oraz własności niskopoziomowe (zależnie od typu danych: kształt, rozmiar, tekstura, kolor, lokalizacja, trajektoria ruchu, tempo, melodia). Opisy typowo obejmują również informacje uzupełniające na temat formy (format, rozmiar, rozdzielczość), dostępności materiału (cena, ochrona praw autorskich), klasyfikacji treści (kategoria tematyczna, kategoria wiekowa odbiorców) i kontekstu (czas i miejsce nagrania, wydarzenie).

## MPEG-7 – Przykład



The diagram illustrates the structure of an MPEG-7 XML document. On the left, a vertical list of five numbered circles (1-5) corresponds to specific parts of the XML code. On the right, a small image of a soccer goal is shown with a yellow rectangular region highlighted. Arrows indicate the mapping from the XML elements to these visual elements.

```

<Mpeg7 ...>
  <Description xsi:type="ContentEntityType">
  <MultimediaContent xsi:type="ImageType">
    <Image id="a102">
      <MediaLocator><MediaUri>kuszczyk.jpg</MediaUri></MediaLocator>
      <TextAnnotation><FreeTextAnnotation>
        Polska - Kolumbia (1:2). Polski bramkarz Tomasz
        Kuszczyk wpuszcza gola na 0:2 po strzale bramkarza Kolumbii.
      </FreeTextAnnotation></TextAnnotation>
      <SpatialDecomposition>
      <StillRegion>
        <TextAnnotation><FreeTextAnnotation>
          Polski bramkarz Tomasz Kuszczyk w wysoku
          myśli że piłka przechodzi nad poprzeczką.
        </FreeTextAnnotation></TextAnnotation>
        <SpatialLocator>
          <Polygon><Coords dim="2 4"> 45 20 0 -20 25 0 60 0
          </Coords></Polygon>
        </SpatialLocator>
      </StillRegion>
    </Image>
  </MultimediaContent>
  ...
</Mpeg7>
  
```

Multimedialne bazy danych (21)

Na slajdzie pokazano przykład opisu MPEG-7 dla obrazka przedstawiającego scenę z meczu piłki nożnej. (Aby przykład zmieścił się na slajdzie niektóre fragmenty zostały zastąpione wielokropkiem.). Znaczenie oznaczonych fragmentów dokumentu jest następujące:

1. Informacja o tym, że opisywany obiekt multimedialny jest obrazem. Następnie początek elementu z opisem obrazu, a w nim jako pierwszy, element zawierający informację o źródle (lokalizacji) obrazu.
2. Ogólny opis treści obrazu w formie adnotacji swobodnym tekstem.
3. Dekompozycja przestrzenna obrazu, umożliwiająca szczegółowy opis poszczególnych jego regionów. W tym wypadku wyróżniony i opisany jest jeden region (zaznaczony na obrazku żółtym obramowaniem).
4. Opis zawartości regionu obrazu w formie adnotacji swobodnym tekstem.
5. Informacje o lokalizacji opisywanego regionu na obrazie. W tym wypadku jest on zdefiniowany jako wielokąt o czterech wierzchołkach. Pierwsze cztery liczby opisują współrzędne X wierzchołków, a kolejne cztery – współrzędne Y. Współrzędne pierwszego wierzchołka są określone w sposób bezwzględny, pozostałych – inkrementalnie względem poprzedzającego wierzchołka.



## Zapytania do multimedialnych baz danych

- Zapytania o zawartość i metadane ją opisujące
  - mechanizm content-based retrieval
- Duża rola interfejsów wizualnych
- Podejście Information Retrieval (IR):
  - wagi przypisywane kryteriom selekcji
  - wyszukiwanie jako proces iteracyjny
  - tolerancja dla niekompletnych wyników

Multimedialne bazy danych (22)

Zapytania do multimedialnych baz danych odwołują się do metadanych zewnętrznych opisujących obiekty multimedialne i do rzeczywistej ich zawartości, reprezentowanej przez metadane semantyczne i sygnałowe. Kluczowe znaczenie mają algorytmy wyszukiwania w oparciu o zawartość (ang. content-based retrieval) operujące na metadanych sygnałowych.

Ważną rolę w przeszukiwaniu multimedialnych baz danych pełnią interfejsy wizualne, za pomocą których zapytania tekstowe są uzupełnione lub zastąpione przez wybranie, dostarczenie lub naszkicowanie wzorca do przeszukiwania kolekcji danych.

Charakterystyczne dla zapytań do multimedialnych baz danych jest podejście w stylu Information Retrieval (IR). Po pierwsze, użytkownik powinien mieć możliwość przypisywania wag poszczególnym kryteriom selekcji np. odwołującym się do koloru i tekstury. Po drugie, zakłada się, że wyszukiwanie ma charakter procesu iteracyjnego, w którym użytkownik modyfikuje parametry zapytania aż do uzyskania satysfakcjonującego rezultatu. Jest to uzasadnione, gdyż np. przy wyszukiwaniu obiektów podobnych do zadanego, trudno jest za pierwszym razem dobrać odpowiednie wagi i próg dla miary podobieństwa. Po trzecie, tolerowane są niekompletne wyniki, w których nie znalazły się niektóre obiekty spełniające zadane kryteria, a być może pojawiły się obiekty tych kryteriów nie spełniające. Takie podejście jest znacząco odmienne od klasycznych zapytań do baz danych, ale w sytuacji gdy zapytania odwołują się niekiedy do subiektywnego odczucia podobieństwa, trudno o obiektywną kompletność wyników wyszukiwania.



## Wyszukiwanie w oparciu o zawartość

- Wyszukiwanie w oparciu o automatycznie wyekstrahowane metadane sygnałowe
- Technologia najlepiej dopracowana dla obrazów
- Alternatywa lub uzupełnienie (!) dla wyszukiwania w oparciu o alfanumeryczne opisy wprowadzone „ręcznie”
- Zaleta: opis słowny czasochłonny i często trudny
- Wiele praktycznych zastosowań
  - wyszukiwanie poprzez podanie przykładu
  - wykrywanie plagiatów

Multimedialne bazy danych (23)

Wyszukiwanie w oparciu o zawartość (ang. content-based retrieval) polega na wyszukiwaniu w oparciu o automatycznie wyekstrahowane metadane sygnałowe. Są to metadane reprezentujące takie właściwości jak:

(a) dla obrazów: średni kolor, udział i lokalizacja kolorów, tekstura;

(b) dla danych audio: melodia, rytm;

(c) dla danych wideo: dynamika ruchu, właściwości obrazów stanowiących klatki filmu.

Wyszukiwanie w oparciu o zawartość jest najlepiej dopracowane dla obrazów. Tylko dla obrazów jest ono przedmiotem standardu SQL/MM. Podobnie, tylko w zakresie przeszukiwania kolekcji obrazów ten mechanizm wspierają dostępne obecnie (wiosna 2006) systemy zarządzania bazą danych ogólnego przeznaczenia (Oracle, IBM DB2 i Informix).

Wyszukiwanie w oparciu o zawartość może stanowić alternatywę lub uzupełnienie dla wyszukiwania w oparciu o alfanumeryczne opisy wprowadzone „ręcznie”.

Zaletą wyszukiwania w oparciu o zawartość jest to, że tworzenie alfanumerycznych opisów zawartości danych multimedialnych jest czasochłonne, a na razie nie może być wykonane bez udziału człowieka. Ponadto, nie wszystkie właściwości można w sposób naturalny opisać słownie, np. tekstury i kształty na obrazach.

Wyszukiwanie w oparciu o zawartość ma szereg zastosowań. Przede wszystkim są to wszelkie interfejsy do zapytań poprzez przykład, a z zastosowań ściśle praktycznych - wykrywanie plagiatów utworów muzycznych, zastrzeżonych znaków towarowych, itp.



## Przykłady zapytań

- Odwołujące się do metadanych zewnętrznych:
  - utwory wykonawcy X dostępne w formacie MP3
  - komedie reżysera X o czasie trwania poniżej 1,5h
- Odwołujące się do znaczenia (metadane semantyczne):
  - przemówienia na temat bezrobocia
  - fotografie przedstawiające prezydenta z premierem
- Odwołujące się do zawartości (metadane sygnałowe):
  - obrazy/utwory podobne do narysowanego/zanuczonego
  - filmy z aktorem przedstawionym na zdjęciu
  - obrazy, w których dominuje kolor czerwony

Multimedialne bazy danych (24)

Slajd przedstawia przykładowe zapytania do multimedialnych baz danych, pogrupowane wg typu metadanych, do których się odwołują.

Przykłady zapytań odwołujących się do metadanych zewnętrznych to:

- „Znajdź wszystkie utwory wykonawcy X dostępne w formacie MP3”;
- „Znajdź wszystkie komedie reżysera X o czasie trwania poniżej 1,5h”.

Przykłady zapytań odwołujących się do znaczenia, realizowanych w oparciu o metadane semantyczne to:

- „Znajdź wszystkie przemówienia na temat bezrobocia”;
- „Znajdź wszystkie fotografie przedstawiające prezydenta z premierem”.

Przykłady zapytań odwołujących się do zawartości, realizowanych w oparciu o metadane sygnałowe to:

- „Znajdź wszystkie obrazy podobne do narysowanego (naskicowanego)”;
- „Znajdź wszystkie utwory podobne do zanuczonego do mikrofonu”;
- „Znajdź wszystkie filmy z aktorem przedstawionym na zdjęciu”;
- „Znajdź wszystkie obrazy, w których dominuje kolor czerwony”.

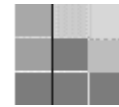
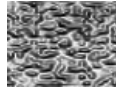
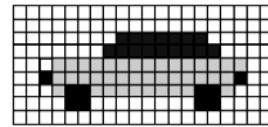
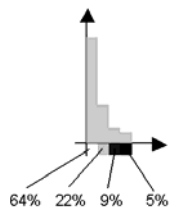
Oczywiście możliwe są też zapytania odwołujące się jednocześnie do różnych typów metadanych, np. „Znajdź wszystkie fotografie w formacie JPEG, w rozdzielczości co najmniej 1024 x 768 pikseli, przedstawiające krajobraz wysokogórski o udziale i lokalizacji kolorów zbliżonej do wskazanego, przykładowego obrazka”.





## Metadane sygnałowe dla obrazów

- Średni kolor
- Histogram kolorów
- Kontury segmentów obrazu (kształty)
- Tekstury
- Lokalizacja kolorów, kształtów i tekstur na obrazie



Multimedialne bazy danych (25)

Jak wspomniano wcześniej, technika wyszukiwania w oparciu o zawartość w multimedialnych bazach danych jest w chwili obecnej najlepiej dopracowana i dostępna dla obrazów. Typowe właściwości obrazów wykorzystywane przy wyszukiwaniu obrazów w oparciu o zawartość to: Średni kolor – wyznaczany w oparciu o podział obrazu na  $n$  próbek i niezależne uśrednianie komponentów RGB koloru;

Histogram kolorów – reprezentujący udział procentowy kolorów w obrazie, najczęściej wyznaczany po zredukowaniu liczby kolorów np. do 256;

Kształty - odkrywane w procesie tzw. segmentacji obrazu, gdzie segmenty są rozumiane jako spójne regiony o tym samym kolorze;

Tekstury - opisujące chropowatość, kontrast, kierunkowość wzorców wypełniających kształty i tło obrazu; reprezentowane w postaci wektorów czy macierzy współczynników liczbowych, wyznaczanych specjalistycznymi algorytmami;

Lokalizacja kolorów, kształtów i tekstur na obrazie – wyznaczana poprzez podział obrazu na siatkę regionów i następnie wyznaczenie dominujących kolorów, kształtów i tekstur dla każdego regionu.



### Interfejsy do wyszukiwania obrazów w oparciu o zawartość

- Najczęściej wyszukiwanie poprzez podanie przykładu
  - wybranego z predefiniowanego zbioru
  - dostarczonego przez użytkownika
  - naszkicowanego za pomocą edytora graficznego
- Możliwość wyboru właściwości i przypisania wag
- Przykład (<http://www.hermitagemuseum.org/>):



Multimedialne bazy danych (26)

Kryteria wyszukiwania obrazów ze względu na zawartość typowo są podawane w formie przykładowego obrazka. Wyszukiwanie sprowadza się do znalezienia obrazów podobnych do podanego przykładu. Często użytkownik ma możliwość wskazania, które właściwości wizualne mają być uwzględnione w testach podobieństwa z opcją przypisania im wag.

Dostarczenie systemowi wzorcowego obrazu może polegać na:

- (a) wyborze z predefiniowanego zbioru obrazków reprezentujących poszczególne kategorie,
- (b) załadowaniu do systemu obrazka dostarczonego przez użytkownika,
- (c) naszkicowaniu obrazu za pomocą prostego, wbudowanego w interfejs edytora graficznego. Edytor ten może być zorientowany na konkretną dziedzinę zastosowań i udostępniać pewne predefiniowane kształty.

Przykładem istniejącego, ogólnodostępnego serwisu umożliwiającego wyszukiwanie obrazów ze względu na zawartość jest serwis internetowy rosyjskiego muzeum Ermitaż (<http://www.hermitagemuseum.org/>). Serwis udostępnia dwie wyszukiwarki oparte o technologię QBIC firmy IBM, która jest obecnie częścią funkcjonalności systemu zarządzania bazą danych IBM DB2:

- (a) QBIC Color Search (<http://www.hermitagemuseum.org/cgi-bin/db2www/qbicColor.mac/qbic?selLang=English>) umożliwiającą wyszukanie obrazów o zadanym udziale kolorów,
- (b) QBIC Layout Search (<http://www.hermitagemuseum.org/cgi-bin/db2www/qbicLayout.mac/qbic?selLang=English>) umożliwiającą wyszukanie obrazów o wskazanej lokalizacji kolorów.

Na slajdzie pokazany jest przykład zapytania o lokalizację kolorów przygotowanego za pomocą interfejsu QBIC Layout Search i przykładowy obrazek z kolekcji obrazów zwróconych przez system dla tego zapytania.



## Ekstrakcja i indeksowanie właściwości

- Ekstrakcja:
  - a priori
  - dynamiczna
  - a priori dla części właściwości, dynamiczna dla innych
- Indeksowanie:
  - właściwości reprezentowane jako punkty w przestrzeni wielowymiarowej
  - wykorzystywane struktury indeksów wielowymiarowych:
    - R-drzewa, pliki kratowe, k-d-B-drzewa, itp.
  - zapytania punktowe, przedziałowe, najbliższy sąsiad

Multimedialne bazy danych (27)

Ekstrakcja właściwości wykorzystywanych do wyszukiwania w oparciu o zawartość z obiektów multimedialnych może odbywać się: a priori – w momencie ładowania obiektów do bazy danych lub dynamicznie – w momencie gdy pojawia się zapytanie, które się do nich odwołuje.

Zaletą ekstrakcji a priori jest większa szybkość realizacji zapytań, gdyż właściwości dla obiektów w bazie danych są już wyznaczone, a być może również poindeksowane. W przypadku ekstrakcji a priori wydłuża się jednak czas ładowania obiektów do bazy danych i rosną wymagania na pamięć dyskową. Z kolei ekstrakcja dynamiczna prowadzi do spowolnienia zapytań, gdyż algorytmy ekstrakcji właściwości są czasochłonne.

Ogólnie, lepszym rozwiązaniem jest ekstrakcja a priori, gdyż pamięć dyskowa nie jest dziś droga, a ważniejsza jest redukcja czasu zapytań (wykonywanych wielokrotnie) niż ładowania danych do bazy (wykonywanego raz). Ekstrakcja a priori daje też możliwość poindeksowania właściwości, co dodatkowo może przyspieszyć realizację zapytań. Poleganie na samej tylko ekstrakcji a priori wiąże się jednak z koniecznością wskazania, które właściwości będą mogły być wykorzystywane w zapytaniach, już na etapie projektowania struktur danych. Dlatego też, możliwym kompromisem jest ekstrakcja a priori dla najważniejszych w danej aplikacji właściwości, z możliwością ekstrakcji dynamicznej tych rzadziej wykorzystywanych.

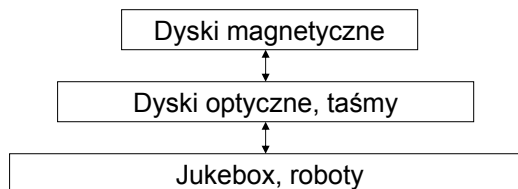
Wyszukiwanie w oparciu o zawartość może być wspierane za pomocą struktur indeksowych. Właściwości obiektów multimedialnych mogą być reprezentowane jako punkty w przestrzeni wielowymiarowej. Z tego powodu, multimedialne bazy danych wykorzystują struktury indeksów wielowymiarowych, takie jak R-drzewa, pliki kratowe, k-d-B-drzewa, itp. Od indeksów wymaga się wspierania zapytań:

- (a) punktowych – do wyszukiwania obiektów o dokładnie takich samych właściwościach;
- (b) przedziałowych - do wyszukiwania obiektów o właściwościach zbliżonych;
- (c) typu najbliższy sąsiad – do wyszukiwania obiektu najbardziej podobnego do danego.

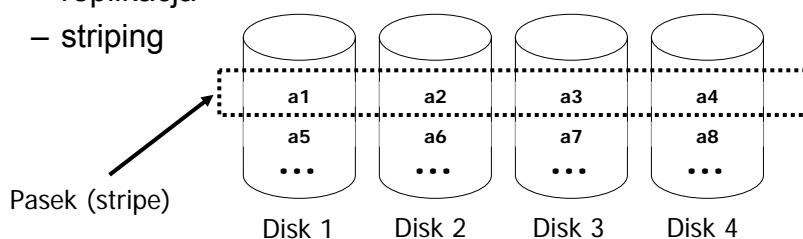


## Składowanie danych multimedialnych

- Hierarchiczne struktury składowania



- Sposoby organizacji danych na nośnikach
  - replikacja
  - striping



Multimedialne bazy danych (28)

Składowanie danych stanowi istotny problem dla danych multimedialnych o szczególnie dużym rozmiarze, czyli przede wszystkim filmów wideo. Pierwszym ograniczeniem jakie może napotkać system składowania danych wideo jest ograniczona pamięć dyskowa. Tradycyjnie problem ten był rozwiązywany przez wykorzystanie hierarchicznych struktur składowania. Idea polegała na przechowywaniu najczęściej żądanych filmów na dyskach magnetycznych (dostępnych on-line), a wszystkich filmów na nośnikach typu dyski optyczne czy taśmy (dostępnych off-line), udostępnianych przez urządzenia typu jukebox albo archiwu obsługiwane przez roboty. Dane zawsze były serwowane z pamięci dyskowej, co pociągało za sobą konieczność migracji danych z pamięci off-line do on-line w przypadku żądania mało popularnego filmu.

Drugim problemem, który miał znaczenie przede wszystkim w przeszłości, była zbyt mała szybkość transmisji danych z dysków. Rozwiązania tego problemu to replikacja i striping. Replikacja polega na duplikowaniu zawartości poszczególnych dysków, dzięki czemu kolejne fragmenty filmu mogą być równolegle odczytywane z kolejnych dysków. Striping to technika umożliwiająca osiągnięcie tego samego efektu, ale bez duplikowania danych i marnowania przestrzeni dyskowej. W tym wypadku dane rozmieszczane są na kolejnych dyskach w formie tzw. pasków (ang. stripe). Fragmenty jednego paska mogą być odczytywane równolegle, zwiększając przepustowość transmisji danych z macierzy dyskowej maksymalnie tyle razy, ile dysków obejmuje pasek.

Rozwiązania przedstawione na slajdzie w ostatnich latach z pewnością straciły na znaczeniu w związku z ciągłym wzrostem pojemności i szybkości dysków magnetycznych połączonym ze spadkiem ich cen. Niemniej, pojemność macierzy dysków serwera w dalszym ciągu jest mniejsza od praktycznie nieograniczonej pojemności archiwów taśmowych obsługiwanych przez roboty, a szybkość transmisji danych z dysków magnetycznych w dalszym ciągu stanowi wąskie gardło w porównaniu z szybkością odczytu danych z pamięci operacyjnej, a nawet szybkością transmisji danych przez nowoczesne sieci komputerowe.



## Prezentacja danych multimedialnych (1/2)

- Problemy z danymi wrażliwymi na opóźnienia
- Transmisja strumieniowa danych wideo
- Warunki dla pojedynczego strumienia
- Obsługa wielu strumieni
  - balans między liczbą żądań a jakością usług
  - szeregowanie strumieni

Multimedialne bazy danych (29)

Problemy prezentacji danych multimedialnych dotyczą danych wrażliwych na opóźnienia, czyli danych audio i wideo. Szczególnie problemy uwidaczniają się dla danych wideo, ze względu na ich duży rozmiar. W praktyce dla danych wideo stosowana jest transmisja strumieniowa z dwóch podstawowych powodów. Po pierwsze, klient może nie dysponować wolną przestrzenią dyskową umożliwiającą pobranie i składowanie całego filmu po swej stronie przed rozpoczęciem odtwarzania. Po drugie, w przypadku transmisji strumieniowej prawie natychmiast po rozpoczęciu przesyłania danych może rozpocząć się odtwarzanie filmu po stronie klienta, bez konieczności oczekiwania na przesłanie całego filmu.

Dla pojedynczego strumienia musi być spełniony oczywisty i z pozoru łatwy do spełnienia warunek, tj. ilość produkowanych danych musi w każdej chwili przewyższać ilość danych konsumowanych. Sytuacja niestety jest o tyle skomplikowana, że poziom konsumpcji danych może być zmienny w czasie ze względu na kompresję danych, a poziom produkcji danych może być zmienny ze względu na konieczność zmiany ścieżek przy odczycie z dysku, zmienne prędkości transmisji danych dla różnych regionów nośnika, itp.

W przypadku obsługi wielu klientów przez wiele strumieni, system musi zadbać, aby do każdego strumienia w odpowiednim momencie wysłać odpowiednią ilość danych. Przyjmując kolejne żądanie system musi zapewnić, aby nie spowodowało to obniżenia poziomu jakości usług obsługiwanym już użytkownikom. Strumienie obsługiwane są w tzw. rundach – w każdej rundzie do każdego strumienia przekazywana jest odpowiednia ilość danych.



## Prezentacja danych multimedialnych (2/2)

- Optymalizacja obsługi wielu klientów
  - batching
  - bridging
  - piggybacking

Multimedialne bazy danych (30)

Optymalizacja strumieniowej transmisji dla wielu klientów sprowadza się przede wszystkim do redukcji liczby strumieni wykorzystywanych do transmisji tego samego filmu.

Zaproponowano w tym zakresie następujące techniki:

**Batching** – polega na odsunięciu w czasie rozpoczęcia transmisji – być może nadejdą identyczne żądania i grupę żądań obsłuży jeden strumień;

**Bridging** – dla strumieni dotyczących tego samego filmu, nieznacznie przesuniętych w czasie, polega na przechowywaniu danych z okna czasowego między strumieniami w pamięci cache;

**Piggybacking** – dla strumieni dotyczących tego samego filmu, nieznacznie przesuniętych w czasie, polega na przyspieszeniu drugiego filmu (poprzez pominięcie niektórych klatek) lub spowolnieniu pierwszego (poprzez wstawienie klatek interpolowanych) tak aby oba żądania były od pewnego momentu obsługiwane przez jeden strumień (zaobserwowano, że wahania prędkości transmisji poniżej 5% są niezauważalne dla człowieka).



## Podsumowanie

- SZMBD = SZBD obsługujący dane multimedialne
- Dane multimedialne znacząco różnią się od tradycyjnych
- Kluczową rolę w opisie i wyszukiwaniu danych multimedialnych odgrywają metadane
- Typowe w multimedialnych bazach danych jest wyszukiwanie w oparciu o zawartość
- Szczególne wyzwanie dla systemów baz danych stanowią dane wideo ze względu na duży rozmiar i wrażliwość na opóźnienia

Multimedialne bazy danych (31)

Od systemów zarządzania multimedialną bazą danych oczekuje się, aby oferowały one tradycyjną funkcjonalność systemów zarządzania bazą danych dla danych multimedialnych. Problemy w implementacji takich systemów wynikają ze specyfiki danych multimedialnych w zakresie modeli danych, przetwarzania zapytań, fizycznych struktur składowania i metod prezentacji.

Kluczową rolę w opisie i wyszukiwaniu danych multimedialnych odgrywają różnego rodzaju metadane. W związku z tym dla multimedialnych baz danych wskazane są modele danych umożliwiające modelowanie obiektów multimedialnych jako obiektów wiążących zawartość binarną i opisujące ją metadane.

Typowe w multimedialnych bazach danych jest wyszukiwanie w oparciu o zawartość, w praktyce najczęściej realizowane poprzez podanie przykładu i wyszukiwanie obiektów podobnych. Zapytania tego typu odwołują się niskopoziomych właściwości obiektów multimedialnych, automatycznie ekstrahowanych z danych specjalistycznymi algorytmami.

Szczególne wyzwanie dla systemów baz danych stanowią dane wideo, ponieważ podobnie jak audio są wrażliwe na opóźnienia w transmisji, a do tego typowo ich rozmiar jest znacząco większy od innych typów danych. Stwarza to problemy odpowiedniej organizacji danych na nośnikach oraz wymaga transmisji strumieniowej przy prezentacji danych.



## Materiały dodatkowe

- V.S. Subrahmanian, Principles of Multimedia Database Systems, Morgan Kaufmann
- B. Thuraisingham, Managing and Mining Multimedia Databases, CRC Press
- H. Kosch, Distributed Multimedia Database Technologies Supported by MPEG-7 and MPEG-21, CRC Press
- J. M. Martínez, MPEG-7 Overview
- J. Yiu-bun Lee, Distributed Video Systems