

Technologie Multimedialne

MULTIMEDIALNE BAZY DANYCH

prezentuje: Paweł Spaleniak

Plan wykładu

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

Multimedia - definicje

- **Obiekt medialny** – informacja zakodowana w jednym medium – media dyskretne: tekst, obraz – media ciągłe: audio, wideo
- **Obiekt multimedialny** – informacja 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

Typy multimediiów

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

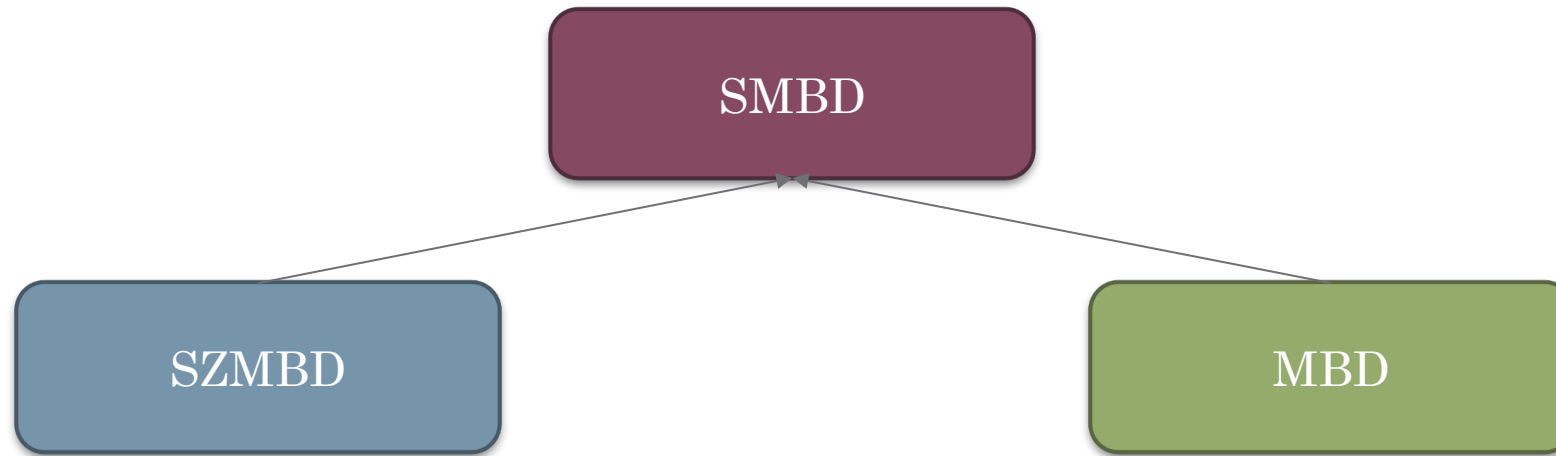
Dlaczego multimedia?

- Otaczająca nas rzeczywistość jest postrzegana jako multimedia
- Dostępność mocy obliczeniowej
- Dostępność urządzeń do składowania danych o dużej pojemności
- Szybkie sieci komputerowe
- Wydajne standardy kodowania i transmisji
- Rozwój technologii mobilnych

Systemy multimedialne w praktyce

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

Systemy multimedialnych baz danych



- **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

Systemy zarządzania multimedialną bazą danych (SZMBD)

- **SZMBD to:**
 - System zarządzania bazą danych
 - 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

Specyfika SMBD

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

Systemy rozproszonych multimedialnych baz danych

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

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

Metadane - przykład



Fot. Shutterstock

Metadane zewnętrzne:

- Autor: Leonardo da Vinci
- Tytuł: Ostatnia wieczerza
- Format: jpeg
- Rozmiar: 1633 x 920

Metadane sygnałowe:

- Średni kolor
- Histogram kolorów
- Kontury segmentów obrazu
- Tekstury

Metadane semantyczne:

Trzynastu mężczyzn (?) siedzących przy stole, najprawdopodobniej rozpoczynających ucztę; w tle widoczne trzy otwory okienne

Metadane obejmują

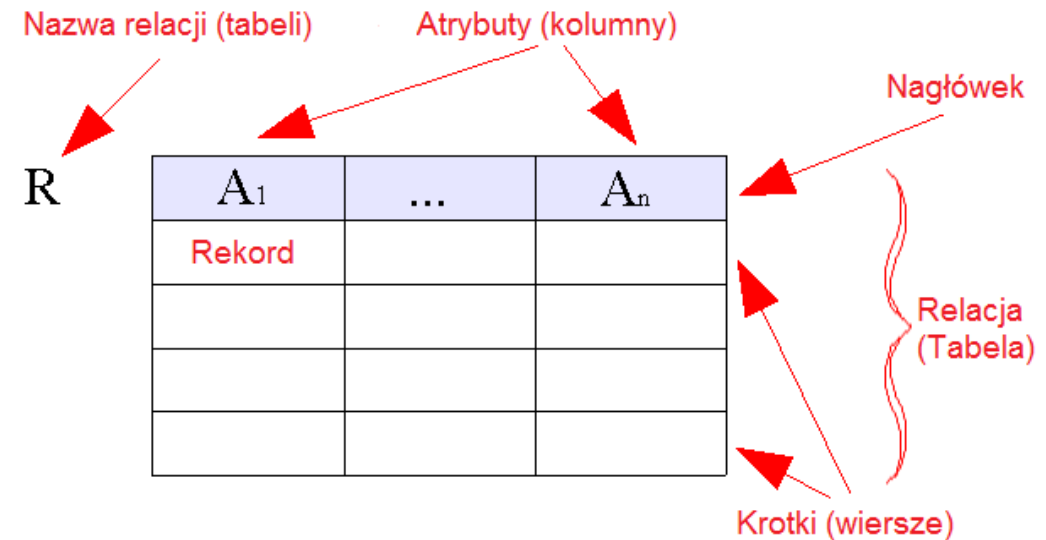
- **ogólne informacje** o obiekcie (tytuł, autor)
- parametry umożliwiające **identyfikację zawartości** (używane przy wyszukiwaniu przy wyszukiwaniu danych),
- **słowa kluczowe** umożliwiające wyszukiwanie zawartości z danej grupy (np. rock lata 90 rock lata 90.),
- informacje o **sposobie prezentacji** (np. wymagana przepustowość łącza),
- właściwości, np. dotyczące **praw autorskich** (zabezpieczenia typu znak wodny, itp.).

Modele danych – podział

- **Model relacyjny** – nieodpowiedni – problemy z reprezentacją złożonych metadanych i wiązaniem ich z zawartością binarną, problematyczne skalowanie horyzontalne
- **Model obiektowy** – odpowiedni, elastyczny, ale nie odniósł sukcesu
- **Model obiektowo-relacyjny** – odpowiedni, praktyczny kompromis
- **Model semistrukturalny** (XML) – odpowiedni do reprezentacji metadanych
- **NoSQL** – nie do końca model ale dobrze wiedzieć ☺

Model relacyjny

- Dane przedstawione są w **postaci relacyjnej**
- Dane grupowane są **w relacje**, które reprezentowane są przez **tabele**
- Nagłówek tabeli - **zbiór atrybutów**
- Każda **krotka** (wiersz) wyznacza zależność pomiędzy danymi w poszczególnych komórkach
- **Klucz główny** – unikalny identyfikator w relacji (może być kombinacją kilku kolumn)
- **Klucz obcy** – zbiór atrybutów jednej relacji wskazujący wartości klucza kandydującego innej relacji
- Operacje na danych, umieszczonych w tabelach dokonywane są za pomocą **języka zapytań (SQL)**

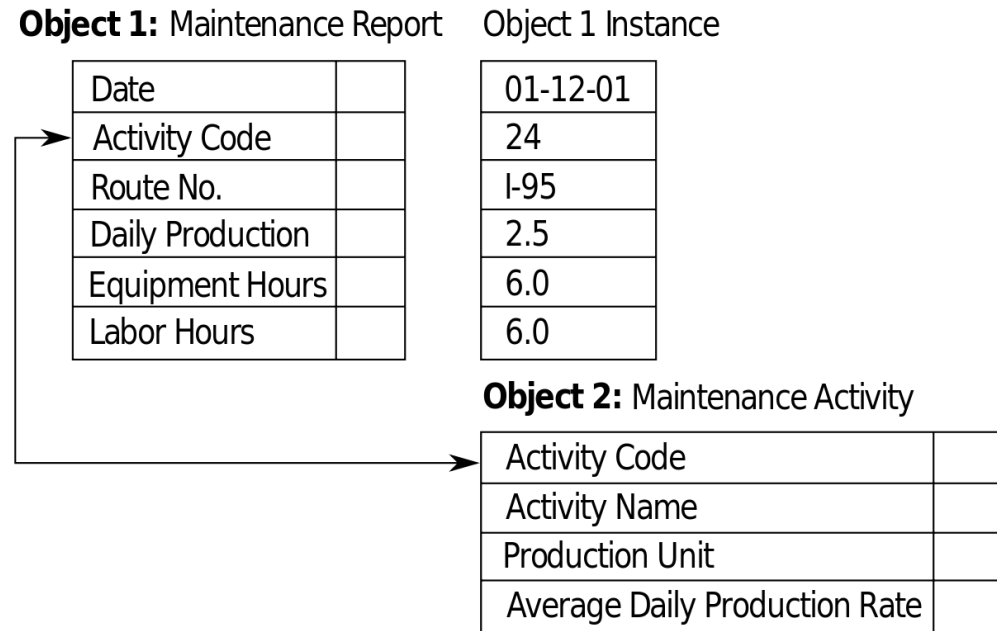


Źródło: Wikipedia

Model obiektowy

- Przystosowany do pracy ze skomplikowanymi **obiektami** danych
- Odzwierciedla **strukturę klas** języków programowania zorientowanych obiektowo
- Obiekty są przechowywane w bazie wraz ze swoimi wszystkimi właściwościami (*properties*)
- Przykład – **JDO (Java Data Objects)** stworzony przez Sun Microsystems, **SQL/MM**

Object-Oriented Model



Źródło: Wikipedia

Model obiektowy

- Obiektowy model bazy danych pozwala użytkownikowi pracować na **wyższym poziomie abstrakcji**;
- Model danych jest **bliższy rzeczywistej strukturze danych** multimedialnych;
- Pozwala tworzyć bazy danych **znacznie szybciej**;
- Hermetyzacja oraz wstawianie inteligencji do zamkniętych obiektów nie zawsze sprawdza się w systemach bazodanowych;
- **Problem wykonania niektórych zapytań**

**Dlaczego model obiektowy jest
lepszy niż relacyjny?**

Oczywiście dla multimedialnych baz danych

Model obiektowo-relacyjny

- **Obsługa obiektowych typów danych w rozszerzonym relacyjnym modelu**
- Intuicyjność i bogactwo środków modelu obiektowego
- Zachowuje najważniejsze zalety modelu relacyjnego (tworzenie zapytań, manipulowanie danymi)

NoSQL

Zalety baz danych NoSQL

- Przewaga NoSQL nad SQL polega na tym, że:
- są bardziej przystosowane i wydajniejsze przy przetwarzaniu Big Data,
- modele danych – brak predefiniowanych schema (czyli schematów tabel, typów danych itp.) powoduje ich większą elastyczność,
- potrafią przetwarzać dane niestrukturalne,
- tańsze i prostsze w utrzymaniu (szczególnie w przypadku prostych baz klucz-wartość) nie wymagają skomplikowanych RDBMS,
- z natury skalowalne (łatwe skalowanie horyzontalne).

NoSQL

Wady baz danych NoSQL

- dla danych, w których występują relacje zalecane są nadal standardowe RDBMS,
- normalizacja w bazach relacyjnych pozwala na uniknięciu redundancji,
- relacyjne bazy danych rokują większą integralności i spójność danych,
- SQL jest szeroko znany,
- brak mechanizmów transakcyjnych – może nie spełniać zasady ACID (ang. Atomicity, Consistency, Isolation, Durability)

Modele danych – systemy i standardy

- **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

Wyszukiwanie danych multimedialnych

- Można wyróżnić 3 poziomy abstrakcji wyszukiwanej informacji
 - Poziom 1 (niski) – pobieranie prymitywnych informacji (dane sygnałowe) w stylu jasność, kolor, kształt, położenie...
 - Poziom 2 (średni) – pobieranie informacji o konkretnych obiektach, używając ich właściwej nazwy (na podstawie wnioskowanych, logicznych własności – „podaj sceny gdzie startuje samolot”)
 - Poziom 3 (wysoki) – pobieranie danych abstrakcyjnych (na podstawie uogólnień) – „znajdź obrazy twarzy wyrażające smutek”

Wyszukiwanie danych multimedialnych

- Sposoby wyszukiwania danych również można podzielić na 3 grupy:
 - **Oparte na atrybutach** (ABR – Attribute-based retrieval)
 - **Oparte na opisie tekstowym** (TBR – Text-based retrieval)
 - **Oparte na rzeczywistej zawartości** (CBR – Content-based retrieval)
- **ABR** – wykorzystuje atrybuty danych, mało efektywne, działanie zbliżone do typowej relacyjnej bazy danych;
- **TBR** – wykorzystuje tekst opisu danych, może mieć określoną strukturę, czasochłonne w czasie wprowadzania, opis może być niezrozumiały dla innych użytkowników;
- **CBR** – bazuje na właściwościach, które są automatycznie pobierane, najtrudniejsze ale najlepsze rozwiązanie dla systemu dla systemu wyszukującego.

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

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

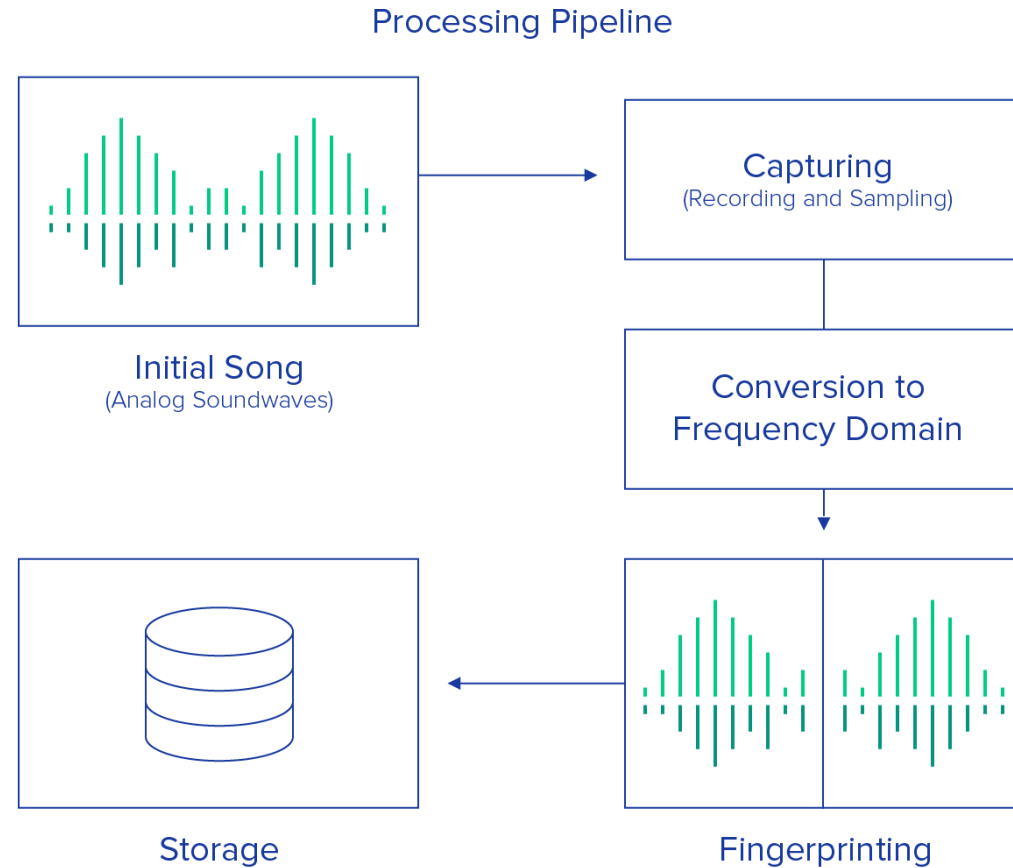
Przykłady zapytań

- **Oparte o metadane zewnętrzne:**
 - 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
- **Oparte na rzeczywistej zawartości (metadane sygnałowe):**
 - obrazy/utwory podobne do narysowanego/zanuczonego
 - filmy z aktorem przedstawionym na zdjęciu
 - obrazy, w których dominuje kolor czerwony

Wyniki wyszukiwania

- Wynik wyszukiwania danych multimedialnych nie jest na ogół dokładny. Prezentowane wyniki powinny być jak najbardziej „bliskie” oczekiwaniom autora zapytania. Dlatego dość istotnym problemem jest zdefiniowanie pewnej miary podobieństwa do wyszukiwanej informacji.
- Jednym z często stosowanych sposobów formułowania zapytania do multimedialnej bazy danych jest zapytanie za pomocą przykładu (QBE – query by example);
- Przykładem wyszukiwania danych multimedialnych wykorzystujących technikę QBE jest system rozpoznawania i wyszukiwania muzyki.
- Przykład może być próbką nagrania z radia lub może zostać zanucony (QBH – query by humming).

SHAZAM – ale jak to działa?



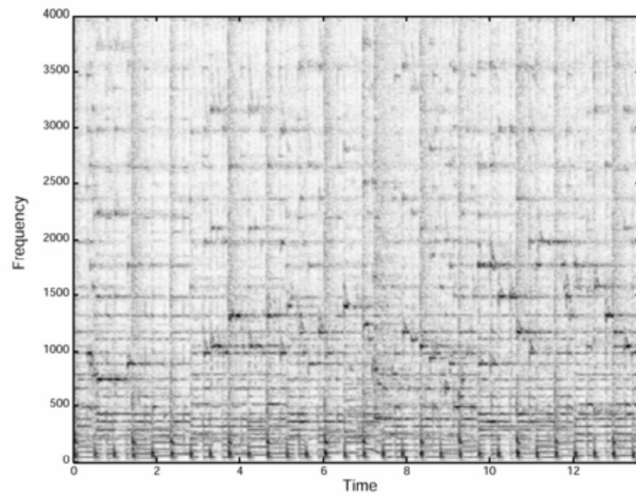


Fig. 1A - Spectrogram

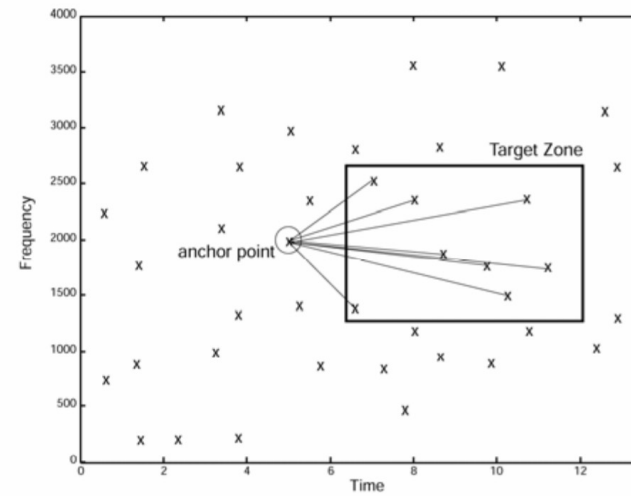


Fig. 1C - Combinatorial Hash Generation

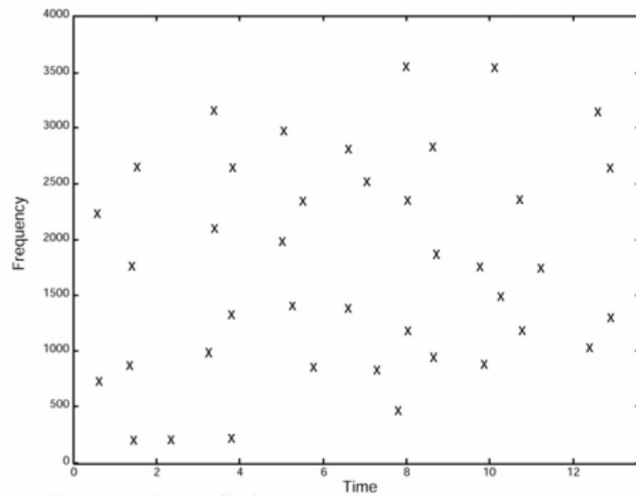


Fig. 1B - Constellation Map

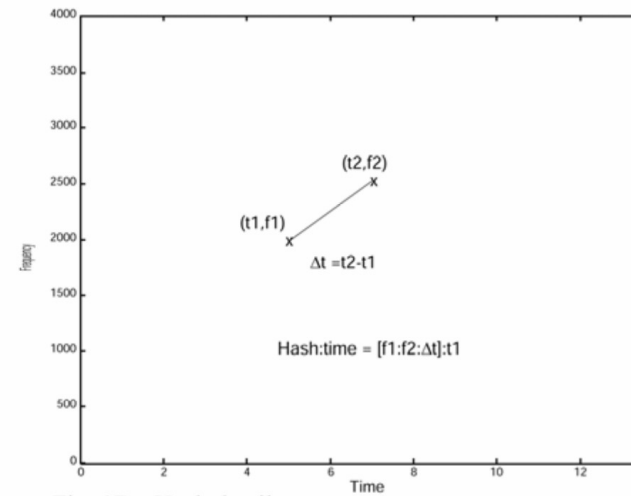
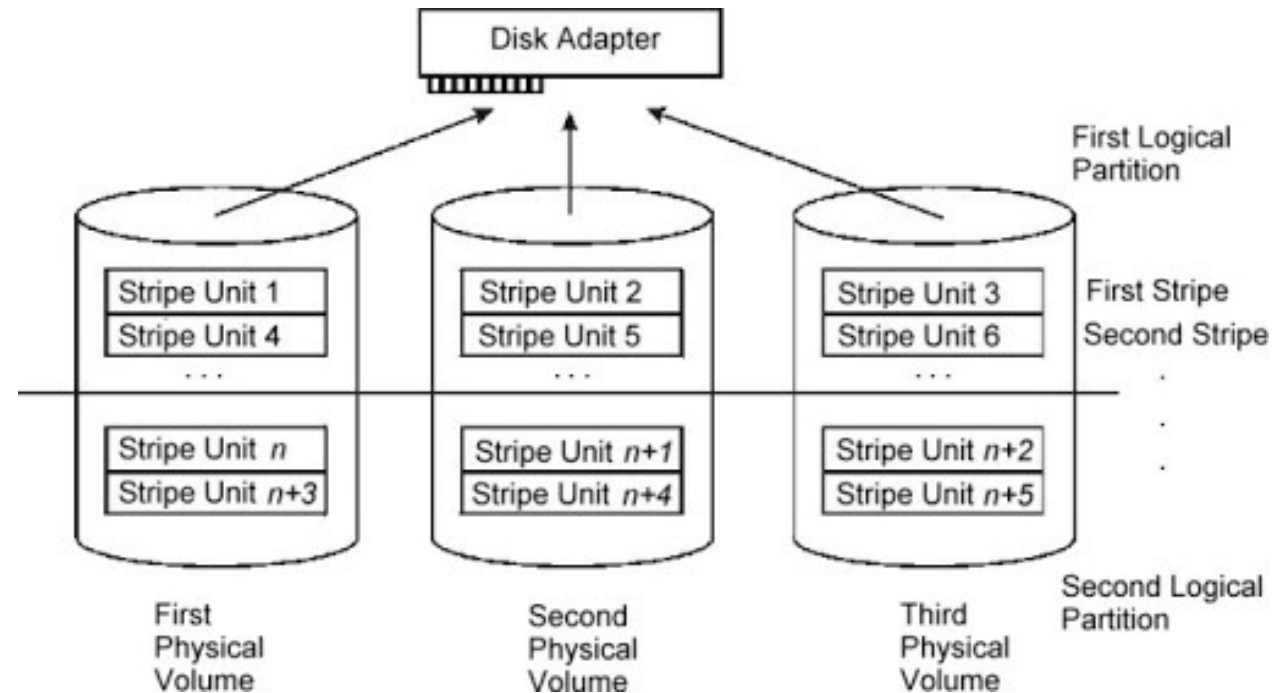


Fig. 1D - Hash details

Źródło: An Industrial-Strength Audio Search Algorithm, Avery Li-Chun Wang

Składowanie danych multimedialnych

- Hierarchiczne struktury składowania
- Sposoby organizacji danych na nośnikach
 - Replikacja
 - Striping



MPEG 7

- Standardy MPEG-1, 2, 4 dotyczyły samej zawarto samej zawartości multimedialnej (kompresja, przesyłanie strumieniowe).
- Standard MPEG -7 dotyczy metadanych opisujących zawartość audiowizualną.
- **Główne składniki standardu MPEG-7:**
 - deskryptory (Descriptors, D) – sposób opisu poszczególnych cech (elementów opisu poszczególnych cech (elementów metadanych),
 - schematy opisu (schematy opisu (Description Schemes Description Schemes, DS) – sposób opisu relacji mi sposób opisu relacji między deskryptorami (również pomiędzy różnymi schematami),
 - język definicji deskryptorów (Description Definition Language, DDL) – język do tworzenia opisów (również do tworzenia nowych schematów lub deskryptorów)
 - schematy klasyfikacji (classification schema, CS) – pojęcia i znaczenia używane do opisu danych
 - narzędzia systemowe (system tools) – przechowywanie i transmisja danych,

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

Bibliografia

- 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