

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

VYHLEDÁVÁNÍ V MULTIMODÁLNÍCH DATABÁZÍCH

SEMESTRÁLNÍ PROJEKT

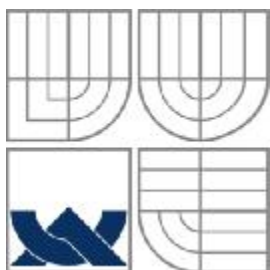
TERM PROJECT

AUTOR PRÁCE

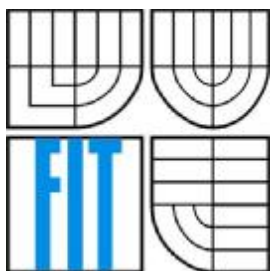
AUTHOR

Bc. Tomáš Krejčíř

BRNO 2008



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

VYHLEDÁVÁNÍ V MULTIMODÁLNÍCH DATABÁZÍCH

MULTIMODAL DATABASE SEARCH

SEMESTRÁLNÍ PROJEKT

TERM PROJECT

AUTOR PRÁCE

AUTHOR

Bc. Tomáš Krejčíř

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Petr Chmelař

BRNO 2008

Vyhledávání v multimodálních databázích

Prohlášení

Prohlašuji, že jsem tento semestrální projekt vypracoval samostatně pod vedením Ing. Petra Chmelaře.

Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Tomáš Krejčř
3.1.2008

Poděkování

Tímto bych rád poděkoval vedoucímu mé práce Ing. Petru Chmelařovi za pomoc při řešení problémů u tvorby této práce.

© Tomáš Krejčř, 2008.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.

Abstrakt

Práce pojednává o vyhledávání informací v multimediálních datech, především pak v audiovizuálních datech a popisuje modely vyhledávání informací mezi které patří hlavně booleovský, vektorový a pravděpodobnostní. Součástí práce je i pojednání o metadatech a podrobněji o standardu MPEG-7 a další důležitou kapitolou je i indexace metadat.

Klíčová slova

Vyhledávání informací, multimédia, databáze, MPEG-7

Abstract

This work deals with information retrieval(IR) from multimedia data, especially from audiovisual data and also describes IR models like Boolean model, Vector space model and Probabilistic model. Part of this work discusses information about metadata and more detailed about standard MPEG-7. Another important chapter is about indexing of metadata.

Keywords

Information Retrieval (IR), multimedia data, database, MPEG-7

Obsah

Obsah.....	1
1 Úvod.....	2
2 Vyhledávání informací.....	3
2.1 Modely IR.....	5
2.2 Vyhledávání.....	8
2.2.1 Podobnostní vyhledávání.....	11
3 Multimediální databázové systémy.....	13
3.1 Multimédia.....	13
3.2 Formáty audiovizuálního obsahu.....	14
3.2.1 Audio.....	14
3.2.2 Video.....	14
3.2.3 Obrázky.....	15
3.2.4 Dokumenty.....	15
3.3 Metadata.....	15
3.3.1 MPEG-7.....	16
3.3.2 MPEG-21.....	19
3.4 Ukládání.....	19
3.5 Indexace.....	19
3.5.1 K-dimensionální strom (kd-strom).....	20
3.5.2 Quad strom.....	21
3.5.3 R-strom.....	22
Závěr.....	24
Literatura.....	25

1 Úvod

Snad vždy byla potřeba nějakým způsobem ukládat či schraňovat nějaké informace, ať už textové nebo obrazové, a vyhledávat v těchto materiálech. Vyhledávání informací neboli Information Retrieval (IR) je od dob vynálezu počítače obor, který jde stále kupředu, ale stále existuje mnoho nevyřešených problémů ať už s uložením dat (především pak multimediálních) nebo se samotným vyhledáváním informací a i přes tyto problémy však v dnešní době existuje množství systémů na vyhledávání informací.

Hlavním cílem této práce bylo seznámit se s problematikou multimediálních databází, především pak databází s audiovizuálními daty, a s problematikou vyhledávání informací z těchto databází.

V následujících kapitolách se pokusím seznámit čtenáře s danou problematikou o kterou jsem se zajímal.

V první části dokumentu se zaměřím na seznámení se s problematikou kolem vyhledávání informací, což zahrnuje popis a vývoj této oblasti a popis jak vypadá typický systém vyhledávání informací. Dále zde jsou popsány modely vyhledávání informací mezi které patří hlavně booleovský model, vektorový model a model pravděpodobnostní. Tyto modely jsou popsány i s jejich funkcemi a vzorci. Posledním popsané téma v této části je vyhledávání v databázích, které se dělí na vyhledávání podle popisu a vyhledávání podle obsahu. Lehce je zde nastíněno jak vypadají procesy v multimediální databázi a dále je zde zmíněno i podobnostní vyhledávání a to především pomocí tříd podobnosti a pomocí Eukleidovské vzdálenosti rysů.

V druhé části se poté zaměřím na pojmy z oblasti multimediální databázové systémy. Z obecného hlediska se zaměřím na typy multimedií a taky formáty audiovizuálního obsahu. Důležitou částí je popis metadat a jejich typy. Zaměřím se poté především na popisná metadata a podrobněji popíši standard MPEG-7. Zmíním se také pár slovy o ukládání dat a na závěr této části popíši indexaci a používané typy jako jsou K-dimensionální strom, Quad strom či R-strom.

Na konci dokumentu je ještě závěrečné zhodnocení této práce a literatura využita ke tvorbě tohoto dokumentu.

2 Vyhledávání informací

Vyhledávání informací (Information Retrieval - IR) je velmi široký termín a základní definice tohoto pojmu je taková, že jde o vyhledávání v nestrukturovaných datech. Mezi nejčastější vyhledávání informací patří vyhledávání v textových dokumentech jako například vyhledávání ve člancích v novinách nebo vyhledávání po internetu. Mezi další typy patří rozpoznávání obrázků, rozpoznávání videa či rozpoznávání hudby. Vyhledávání informací je činnost, jejímž cílem je identifikace relevantních dokumentů nebo informací v informačních zdrojích (např. fulltextových databázích), souvisí s reprezentací, skladováním, organizací a přístupem k informacím.

Rozdíly mezi vyhledáváním dat a vyhledáváním informací jsou zobrazeny v tabulce Tab. 2.1.

	Vyhledávání dat	Vyhledávání informac
Shoda	přesná shoda	částečná shoda, nejlepší shoda
Odvození	dedukce	indukce
Model	deterministický	pravděpodobnostní
Dotazovací jazyk	strukturovaný kompletní	přirozený nekompletní
Potřeba	data	informace
	shodná	relevantní

Tab 2.1 Rozdíly mezi vyhledáváním dat a vyhledáváním informací

Nyní o rozdílech vyhledávání o něco více podrobněji. U vyhledávání dat běžně hledáme naprostou shodu, tedy kontrolujeme zda porovnávaný vzorek je naprosto shodný s porovnávaným. U vyhledávání informací to může být stejné, avšak většinou se pokoušíme najít ty, které se částečně shodují a z nich vybereme několik nejlepších.

Odvození využito u vyhledávání dat je deduktivního typu, tj. aRb a bRc potom aRc . U vyhledávání informací je mnohem běžnější využít indukční hypotézy. Tento rozdíl vede k popisu vyhledávání dat deterministicky, ale vyhledávání informací pravděpodobnostně.

Dotazovací jazyk pro vyhledávání dat je obecně umělého rázu s omezenou syntaxí a slovníkem a u vyhledávání informací se upřednostňuje využití přirozeného jazyka i když existují výjimky.

Množství automatických systémů vyhledávání informací je experimentálních. Takové systémy se provozují především v „laboratorních“ podmínkách. Pochopitelně 2 systémy jsou různě hodnoceny. Systémy vyhledávání informací v reálném světě jsou hodnoceny podle „spokojenosti uživatele“ a ceny, kterou je uživatel ochoten zaplatit. Experimentální systémy jsou hodnoceny porovnáním experimentů se standardy, speciálně vykonstruovány pro daný účel.

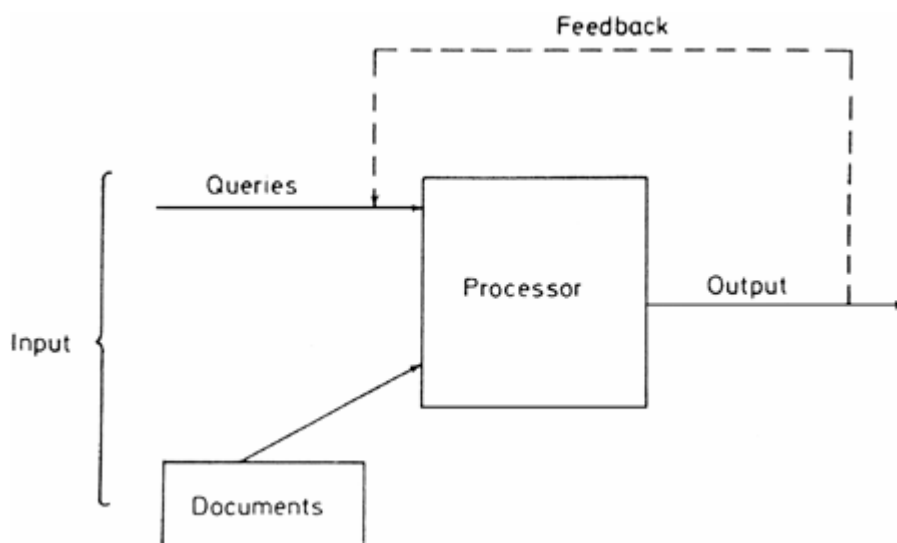
Od 40. let 19. století problém ukládání informací a vyhledávání v nich přitahuje větší a větší pozornost. Je to především tím, že máme obrovské množství informací, ke kterým chceme přesně a rychle přistupovat a to je čím dál tím těžší. S nástupem počítačů začíná doba využívání jich k poskytování rychlých a inteligentních systémů pro vyhledávání. Některé problémy vyhledávání jako například „uskladnění“ informací byly vyřešeny počítači, nicméně problém efektivity vyhledávání zůstal nevyřešen.

V principu je uložení informací a vyhledávání v nich jednoduché. Předpokládá se „sklad“ dokumentů a osoba(uživatel) formulující dotaz, ke kterému je odpověď množina dokumentů, které uspokojí potřeby dotazující osoby. Zjištění hledaných informací lze provést přečtením všech dokumentů ve „skladu“. Toto řešení je však velmi nepraktické. Uživatel nemá čas nebo si spíše nepřeje trávit čas čtením celých dokumentů. Proto když se staly vysoko-rychlostní počítače dostupnými, myšlenky byly takové, že počítač bude schopen „přečíst“ veškeré dokumenty a vyextrahovat jen ty relevantní. To se také brzy stalo použitím textového přirozeného jazyka dokumentu, ale zůstal nevyřešen problém jak charakterizovat obsah dokumentu. Softwarové pokusy o automatickou charakterizaci, tedy o provedení lidské činnosti „čtení“ je vskutku problém. Více specificky, „čtení“ zahrnuje pokusy vyextrahovat informace z textu, a to jak syntaktické tak sémantické, a využití těchto informací k rozhodnutí, zda každý dokument je relevantní nebo není. Obtížné není jenom to jak extrahovat informace, ale také jak tyto informace použít k rozhodnutí relevance. Poměrně pomalý pokrok v moderní jazykovědě a viditelný nedostatek u překladatelských strojů ukazují, že tyto problémy jsou do značné míry nevyřešeny.

Smysl strategie automatického vyhledávání je vyhledat všechny relevantní dokumenty . Pokud charakterizace dokumentu je vyřešena, mělo by to být pokud dokument je relevantní vzhledem k dotazu, což umožní dokumentu být vyhledán v odpovědi na dotaz. Důležitým pojmem je indexování a pokud je indexování provedeno automaticky, tak se předpokládá, že využití stejné automatické analýzy pro text dokumentu nebo dotazu, výstup bude reprezentace obsahu a pokud dokument je relevantní vzhledem k dotazu, výpočetní procedury to ukážou.

Rozumově je pro člověka možné zjistit relevanci dokumentu vzhledem k dotazu. Pro počítač potřebujeme udělat model, který dokáže spočítat rozhodnutí relevance. Je zajímavé všimnout si, že většina výzkumů v oblasti vyhledávání informací může být ukázáno s různými aspekty modelu. Přeloženo z [1].

Na obrázku Obr. 2.1 lze vidět jak vypadá typický **systém vyhledávání informací**. Takový systém se skládá ze 3 hlavních komponent: input(vstup), do kterého patří Queries(dotazy) a Documents(dokumenty), procesor a output(výstup). U on-line systémů je součástí i tzv. Feedback.

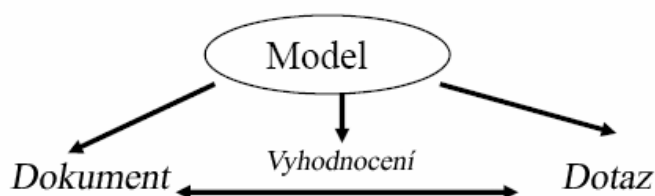


Obr. 2.1 Typický systém vyhledávání informací [1]

Hlavním problémem vstupní „komponenty“ je získat reprezentaci všech dokumentů a dotazů vyhovující pro počítačové zpracování. Reprezentace dokumentu by mohla být extrakce slov, která jsou považována za podstatná. Pokud je systém on-line, je možné, aby uživatel měnil své požadavky během vyhledávání. Taková procedura je běžně zvaná jako „feedback“ (zpětná vazba). Procesor je část systému, který se týká procesu vyhledávání. Proces může vyžadovat strukturování informací vhodným způsobem, jako např. klasifikování. Na závěr jsme se dostali k výstupu, který je obvykle množina citací nebo čísel dokumentů.

2.1 Modely IR

Modely IR vysvětlují strukturu a procesy systémů vyhledávání informací a objasňují jejich obecnost jako protiklad k specifické charakteristice. Model IR dává odpověď pro mechanismus rozhodování relevance. Modely IR odpovídají na otázky relevance dotazu k dokumentům v databázi: „Jaké dokumenty mají být výsledkem dotazu?“ a „Jaké bude jejich uspořádání pro prezentaci uživateli?“.



Obr. 2.2 Schéma Modelu IR

Existuje mnoho modelů IR. Mechanizmy na rozhodnutí relevance mohou být buď přesné nebo flexibilní a reprezentace dat může mít proměnný stupeň abstrakce. Mezi klasické modely patří

booleovský model, vektorové modely či pravděpodobnostní modely. Mezi alternativní modely patří modely neural network, SVM či Fuzzy set.

Klasický **booleovský model** je jeden z nejstarších vůbec a v minulosti býval hojně používán v knihovnických informačních systémech. Na dotaz vrací jako výsledky ty dokumenty, které obsahují slova z dotazu. V základní variantě neumožňuje stanovování relevance a funkce podobnosti je vyčíslována pouze s hodnotami 0 (zásah nenalezen) nebo 1 (zásah). Model předpokládá, že dokument d je popsán množinou klíčových slov – indexační termíny k . Termíny jsou obvykle podstatná jména. Binární rozhodovací kritérium pro tento model je založené na přítomnosti, resp. absenci daného klíčového slova a neuvažuje jeho váhu (pravdivost, dokazatelnost). Všechny dotazy jsou tvořeny termy a logickými spojkami *and*, *or*, *not* a závorkami. Výhody tohoto modelu je efektivnost a využívá se pro dotazování, mezi nevýhody patří to, že výsledkem je příliš (málo) dokumentů a neřeší uspořádání [2].

Zde uvádím několik typů binárních porovnávání:

$$\emptyset \text{ Jednoduché: } M = |X \cap Y|$$

$$\emptyset \text{ Diceovo: } M = \frac{|X \cap Y|}{|X| + |Y|}$$

$$\emptyset \text{ Jaccardovo: } M = \frac{|X \cap Y|}{|X \cup Y|}$$

$$\emptyset \text{ Kosinové: } M = \frac{|X \cap Y|}{\sqrt{|X|} \cdot \sqrt{|Y|}}$$

$$\emptyset \text{ Overlapovo: } M = \frac{|X \cap Y|}{\min(|X|, |Y|)}$$

Dalším modelem je **vektorový model**. Vektorový model umožňuje jemnější výpočet relevance. Myšlenka je opět velice jednoduchá. Každý text či dotaz (prostě jakákoliv skupina slov) je reprezentován bodem v n -rozměrném souřadnicovém systému. Tento bod představuje i vektor (začínající v počátku souřadnic), a tak dostal model i svůj název.

Konstrukce bodů je taková, že čím blíže jsou k sobě, tím více reprezentují podobný (dokonce téměř totožný) text. Dále ale budu hovořit o vektorech, nikoliv bodech. Skalární součin dvou vektorů vychází největší, pokud mají tyto vektory stejný směr. Jako nulový vychází, pokud mají vektory směr opačný. Tohoto jevu bývá s úspěchem využito pro vlastní kalkulaci podobnosti.

Předpokládejme, že v textech máme n rozdílných slov. Toto n také určuje onu n -rozměrnost našeho vektorového systému. Každý vektor pak na souřadnici - odpovídající danému slovu - obsahuje jeho četnost (ať již v jednotlivém dokumentu, nebo třeba dotazu).

Podobnost stanovíme jako skalární součin dvou vektorů - ty budeme pro přehlednost označovat jako $v()$. V případě podobnosti dotazu a dokumentu pak tato formule vypadá takto:

$$\text{sim}_{v(d_j),v(q)} = \text{sum}(i=1..n; w_{i,j} w_{i,q})$$

kde $v(d_j) = (w_{1,j}, \dots, w_{n,j})$ a $v(q) = (w_{1,q}, \dots, w_{n,q})$. Hodnoty $w_{i,j}$ udávají počet slova t_i v j -tém dokumentu. Podobně $w_{i,q}$ udává počet slova t_i v dotazu q .

V pokročilých implementacích již $w_{i,j}$ nepředstavuje četnost, ale naopak důležitost, která má většinou výchozí kalkulaci v četnosti.

Dále nebývá vhodné nechat růst vektory (jejich délku) nade všechny meze. Proto je výhodné normalizovat používané vektory na jednotkovou délku.

Tím je pak do značné míry zajištěno, že původně dva krátké vektory (protože obsahovaly jen malý počet slov), jdoucí týmž směrem, nebudou při výpočtu podobnosti pomocí skalárního součinu přebity dvěma velice dlouhými vektory, které již ale nejdou tak úplně týmž směrem. Tak dostáváme základní formuli:

$$\text{sim}_{v(d_j),v(q)} = \text{sum}(i=1..n; w_{i,j} w_{i,q}) / (\text{length}(v(d_j)) \text{length}(v(q)))$$

Funkce podobnosti (kosinová míra) klasifikuje podobnost vektoru dokumentu d_j k vektoru dotazu q – kosinus odchylky vektorů:

$$\text{SIM}_{\cos}(q, d_j) = \frac{\sum_{k=1}^n q_k * w_{kj}}{\sqrt{(\sum_{k=1}^n q_k)^2 * (\sum_{k=1}^n w_{kj})^2}}$$

Dotazy (rozsahové a na k nejbližších sousedů) jsou realizovány pomocí funkce podobnosti.

Idea **pravděpodobnostního modelu** je taková, že v případě určitého dotazu se podaří výsledné dokumenty uspořádat s klesající pravděpodobností relevance, tak bude efektivita systému nejlepší. Relevance v ID je nejistá a proto je důležitá pravděpodobnost a teorie rozhodování. Nejlepším řešením tohoto modelu je potom Bayesův teorém, který je zobrazen na obrázku Obr. 2.3.

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Obr. 2.3 Bayesův teorém

Důležitým pojmem je **Relevance(významnost, důležitost)**, což je velmi subjektivní pojem. Je to měřítko toho, jak silně nějaká věc ovlivňuje realitu, jak kvalitní je informace nebo jak kvalitní je určitá teorie. Relevance je důležité kritérium kvality informace.

Významnost velmi silně závisí na kontextu. Dává objektům určitou *váhu* v závislosti na vztazích mezi objekty. Čím vyšší je váha objektu, tím více je potřeba stav objektu sledovat, neboť jeho změny významně působí na okolí. To co může být v jednom kontextu relevantní, může být v jiném kontextu zcela nevýznamné.

Určování důležitosti objektů se provádí pomocí určení důležitostí konkrétního subjektu a může být prováděno objektivními nebo subjektivními metodami. Objektivní metody spočívají v tom, že se určí konkrétní parametry a rozsah povolených hodnot. Tyto parametry se následně úplně a jednoznačně zaznamenají. Pokud není možné zjistit jednoznačně a přesně hodnoty parametrů, provádí se vyhodnocení více nebo méně intuitivně. Tzn. že vyhodnocení je založeno na znalostech, zkušenostech, osobních dojmech. Taková metoda je ale subjektivní.

Měření efektivity - U relevance se předpokládá, že dokument je nakonec předurčený, aby byl relevantní nebo ne podle uživatele. Efektivita je pouze míra schopnosti systému uspokojit uživatele v termínech relevance získaného dokumentu. Nejprve se musím soustředit na míry efektivity přesností a úplností.

V této chvíli je užitečné představit známou tabulku „souvislostí“:

	Relevantní	Ne-relevantní
Získaný (true)	$A_y + B_y$	$A_n + B_y$
Nezískaný (flase)	$A_y + B_n$	$A_n + B_n$

Z této tabulky mohou být odvozeny míry efektivity.

$$PRESNOST(PRECISION) = \frac{|A_y \cap B_y|}{|B_y|}$$

$$ÚPLNOST(RECALL) = \frac{|A_y \cap B_y|}{|A_y|}$$

2.2 Vyhledávání

Vyhledávání relevantní informace v audiovizuálních datech je značně náročný, ale elementární úkol každého MMDBMS. Z toho důvodu požadujeme, aby bylo vyhledávání co možná nejefektivnější. Proto je v naprosté většině případů omezeno na metadata (obsahující popis i rysy), která musí být předem extrahována a vhodně indexována, na čemž výrazně závisí efektivita vyhledávání.

Vyhledávací algoritmus je také důležitý, ale především musí umět interpretovat správně daná metadata.

Metadata jsou relativně běžná, strukturovaná data, je tedy možné využít libovolný strukturovaný dotaz. Tak se tomu děje v případě vyhledávání dle textového popisu dat, které je v dnešní době nejběžnější, např. vyhledávání obrázků na internetovém portálu Gogole.

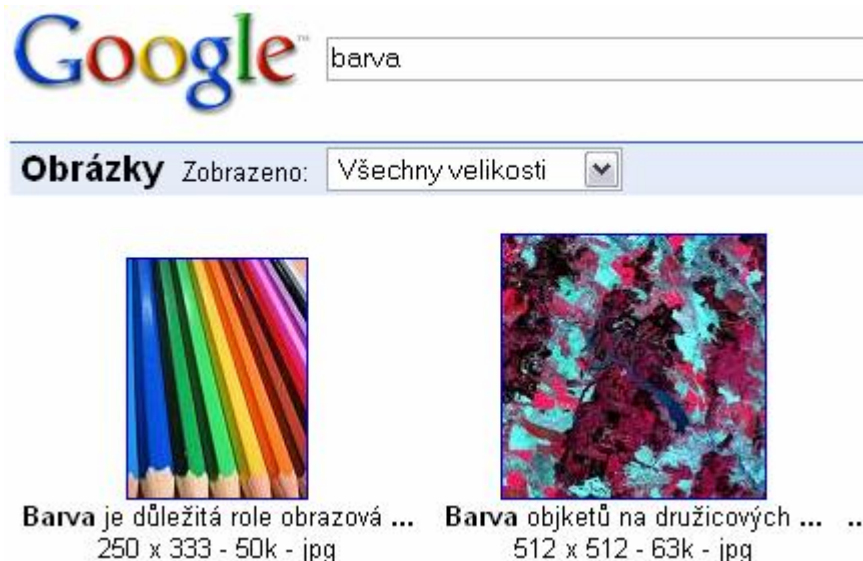
Pro vyhledávání dle obsahu audiovizuálních dat je výše popsáný způsob dosti nepraktický. Například při dotazu na nalezení jisté textury by bylo nutné znát všech 30 koeficientů deskriptoru. Takovýto způsob se nazývá vyhledávání specifikací vlastností.

Některé MMDBMS dokáží specifikaci vlastností výrazně zjednodušit. Je možné například dotaz načrtnout, tj. specifikovat barvu a tvar objektu, nebo zanotovat melodii. Přestože je tento způsob pro lidi velmi intuitivní, netěší se takové podpoře, jako podobnostní vyhledávání, se kterým má mnoho společného [4].

Na základě způsobu popisu rozlišujeme 2 typy vyhledávání:

- Ø **Vyhledávání dle popisu dat** (description-based retrieval). Mezi základní znaky patří to, že se jedná o popis sémantiky dat. Vyhledávání se provádí pouze podle textové anotace, tzn. že je nutné přiřadit název, zapsat klíčová slova atd. a to nelze obecně provést žádným automatizovaným systémem. Popis musí vytvořit člověk při jejich vytváření, takže anotace je vždy určitým způsobem nepřesná, subjektivní, neúplná či jakýmkoliv způsobem zavádějící. Příkladem je vyhledávání obrázků na internetu pomocí portálu Gogole zobrazeno na obrázku Obr. 2.4.

- Ø **Vyhledávání dle obsahu dat** (content-based retrieval) je metoda popisu syntaxe a to často plně automatizovaná což nahrazuje nedostatky této metody, nejvíce pak kvality popisu sémantiky. U popisu obrazových dat máme k dispozici dominantní barvu, textury, tvary, umístění, histogram nebo také např. popis obličeje. Výhodou této metody je skutečnost, že se vyhledává podle skutečného obsahu. Mezi nevýhody patří existence mnoho různých metod modelování struktury a sémantiky obsahu. Otázkou je „kterou vybrat“ ?



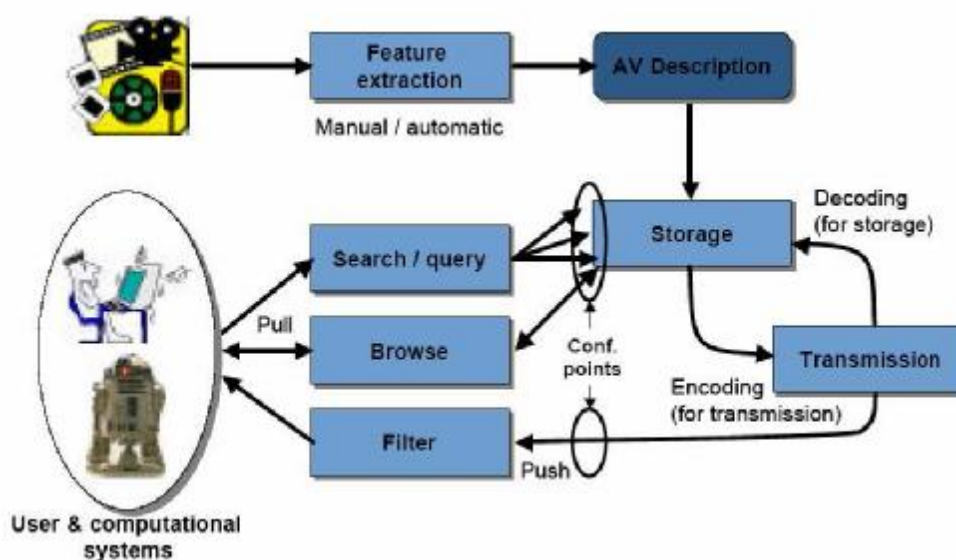
Obr. 2.4 Vyhledávání obrázků na portálu Google

Kombinace manuálního a automatického vyhledávání se dá dobře využít v případech, kdy známe název snímku či nějaký útržek, datum pořízení nebo klíčová slova a z jeho obsahu mohou být extrahovány další vlastnosti jako např. diagnóza.

Ve vyhledávání zaměřeném na obsah se obvykle používají 2 typy dotazů:

- Ø předložení vzorového obrazu (similarity search) – využívá se porovnání deskriptorů extrahovaných z obrázku dotazu s deskriptory uloženými v databázi.
- Ø specifikace vlastností – prvky, ze kterých je tvořen deskriptor jsou porovnány s těmi v databázi.

Procesy v multimediální databázi lze zjednodušeně pochopit z obrázku Obr. 2.5.



Obr. 2.5 Schéma procesů MMDB [9]

První a velmi důležitou součástí procesu je získání multimediálních dat, na kterých je celý proces postaven. Z multimediálního obsahu se získá audiovizuální popis (Feature extraction) a to pomocí manuální nebo poloautomatické extrakce. Tyto extrakce tvoří deskriptor (AV Description), které jsou uloženy spolu s daty do databáze (Storage). Uživatelé či jiné služby mohou poté databázi dotazovat (Search / query) a výsledky těchto dotazů si prohlížet (Browse). Zobrazené výsledky lze filtrovat.

2.2.1 Podobnostní vyhledávání

Člověk umí vyjádřit podobnost jistým „nativním“ způsobem. Podobnost proto musí brát v úvahu především modely lidského vnímání [5].

Princip podobnostního vyhledávání je relativně jednoduchý. Předpokládá, že pro vyhledávání požadované informace jsme schopni dodat nějaký příklad požadovaného audiovizuálního obsahu. Z tohoto média MMDBMS extrahuje stejné rysy, jako by jej chtěl uložit do databáze. Tato metadata dotazu jsou poté využita pro vyhledávání.

Běžně se používají první dvě podobnostní funkce[4]:

- Ø Pomocí tříd podobnosti. Máme-li množinu podobných objektů, pak relace podobnosti na této množině je reflexivní (soběpodobnost) a symetrická (jako dvojčata). Ale objekt této množiny není podobný objektu jiné množiny. Objekty mohou být označeny (tag, label) například ručně.
- Ø Pomocí Eukleidovské vzdálenosti rysů. Máme-li spojitě (číselné) intervaly, a je-li možné mezi nimi měřit vzdálenosti (měla by platit symetrie, nezápornost a trojúhelníková nerovnost), pak můžeme tuto vzdálenost označit jako míru podobnosti. Nicméně záleží pouze na návrhu vektoru rysů, jestli bude tato podobnost odpovídat také jejímu lidskému pojetí.
- Ø Podobnost odvozená od (čas)prostorových dat [6]. Prostorové pro topologické a vzdálenostní dotazy. Temporální dotazy (úsek, kdy zmizel) a jejich kombinace (sledování více kamerami, kde se vyskytoval objekt dříve).
- Ø Transformační vzdálenost. Transformace pracují přímo s obsahem pro porovnání každé dvojice obrázků, což je silně neefektivní. Výhodou je možnost vyjádřit subjektivitu. Jedná se pouze o akademickou práci, realizace mi není známa.

Podobnostní vyhledávání v (časoprostorových) datech má několik úskalí.

- spojitý charakter – objekty je nutné vyhledávat pomocí plovoucího okna
- objekt může být různě velký a znamenat to stejné (portrét, dav) nebo naopak. Proto se používají algoritmy s adaptivní velikostí okna, které jsou invariantní velikostně (adaptive scaling) [7] a časově (dynamic time warping) [8].

3 Multimediální databázové systémy

V dnešní době je obrovská spousta dat a aplikací pracujících s daty, které požadují databáze. Je to především proto, že zaručují konzistenci, integritu, souběžnost, bezpečnost a dostupnost dat.

Multimediální databázi může být např. kolekce obrázků, audia, videa, časových řad, textů, XML, atd. Obecně je to kolekce nestrukturovaných dat, u kterých vnitřní struktura i sémantika je skrytá a nejednoznačná - závislá na aplikaci, datech, i subjektivitě uživatele. U relačních databází se využívají dotazy na shodu což u multimediálních databází nestačí.

V dnešní době jsou multimediální databáze prostě nepostradatelné. Biometrické databáze jsou obrazové databáze a využívají se především u policie na identifikaci osob. Video kolekce se využívají na záznamy z bezpečnostních kamer (letišť, supermarkety, centra měst, atd.) či filmové kolekce a zvukové databáze se využívají např. při zpracování řeči.

3.1 Multimedia

Multimediální data jsou např. zvukové stopy(audio), video záznamy či obrazová data, ale také mezi tyto data patří například strukturované texty či rukou psané poznámky. Můžeme tedy říct, že multimediální data jsou nestrukturovaná data, které dělíme na typ vizuální a audio:

Vizuální data – data související se zrakovým vnímáním

Ø 2D

- o obrazová data (fotky)
- o vektorová grafika – data je možné převést na křivky, plochy či podobně, ale mohou být uchována v podobě bitmapy
- o text – je zvláštní druh vektorové grafiky

Ø 3D (prostorové)

- o modely – př. postava
- o objekty – reálná instance modelu s vlastnostmi získanými z 2D rastrů
- o ne-objekty – určité oblasti obrázku nejsou objekty. Př. obloha, moře

Ø 3-4D (časoprostorové)

- o pohyblivé obrázky – iluze pohybu (video bez zvuků)

Audio data

- Ø hudba – rozlišujeme spec. harmonické vlastnosti zvuků
- Ø řeč
- Ø zvuky nezařazené modelujeme podobně jako obrázky

Pokud pomineme některé speciální případy, pak každý multimediální obsah spadá plně, nebo po částech, do jedné nebo více uvedených kategorií. Například webová stránka obsahuje formátovaný text, (vektorovou) grafiku, rastrové obrázky, ale i videosekvence – pohyblivé obrázky se zvukem (kombinace řeči a hudby). Jejich kombinace se nazývají audiovizuální (AV) data. Převzato z [4].

3.2 Formáty audiovizuálního obsahu

Pro univerzální klienty MMDBMS umožňující prohlížení a tvorbu multimediálních dat, by se měly rozlišovat pouze dva typy – vizuální a audio. Metadata je případně možné reprezentovat jako vizuální data.

3.2.1 Audio

Audio(zvukové) záznamy mohou být v mnoha formátech, mohou mít různé metody komprese či různé vzorkovací frekvence. Takové informace často bývají dostupné jako metadata v hlavičce souboru a dají se využít při vyhledávání.

Zvuk je digitalizovaný audio signál, který je reprezentován diskretními vzorky v časové nebo frekvenční doméně. Známé audio formáty jsou většinou odvozeny od zvukových stop standardních multimediálních kontejnerů i když byly vyvinuty zvlášť, př. formát MPEG-1 Audio Layer III (MP3). Současné MP3 využívají specifikaci z MPEG-2, resp. MPEG-2.5. Vylepšením je zvuk AAC, který využívají MPEG-4. Mezi další formáty patří WMA, Vorbis, AC3 či RealAudio.

3.2.2 Video

Počet formátů video záznamů je mnohem větší než u zvukových záznamů. Existuje mnoho parametrů obrazu, tj. například rozlišení, délka záznamu, barevná hloubka či použité kódování. Podobně jako u audio záznamů i tyto informace lze většinou získat z metadat nebo hlavičky souboru a také podle nich lze vyhledávat.

Video lze popsat jako pohyblivé obrázky se zvukem. Video ale může obsahovat i jiná multimediální data jako např. titulky. Video data obsahují obrovské množství nadbytečné informace. Hlavní a základní idea je tedy odstranit redundanci v rámci video-rámců i mezi nimi. Přední kompresní algoritmy byly vyvinuty v rámci spolupráce skupiny odborníků na pohyblivé obrázky MPEG (noviny picture experts group). Původní formát MPEG-1 využíval pro odstranění redundance jednotlivých snímků algoritmus JPEG. Nástupcem byl MPEG-2 navržen pro digitální TV a rozhlasové vysílání prostřednictvím satelitů. Další nástupce byl MPEG-4, který využívá kódování objektů(video, audio či 3D objekty). Mezi další formáty patří AVI, ASF, QuickTime či WMV.

3.2.3 Obrázky

I obrázky lze ukládat v několika formátech. Jako předcházející data i obrázky obsahují metadata, ze kterých můžeme získat např. rozlišení obrázku, jeho barevnou hloubku, formát obrázku či jiné informace. Také těchto dat se dá využít pro vyhledávání. Obrázky se dělí na bitmapové (raastrové) a vektorové.

Obraz je popsán jednotlivými barevnými hodnotami bodů (pixelů), uspořádaných do mřížky. Mezi zástupce patří JPEG, PNG, GIF, BMP či TIFF. Vektorová grafika je reprezentována geometrickými útvary (přímkami, křivkami, plochami). Patří sem např. PS, CDR, DXF, DVI, aj.

3.2.4 Dokumenty

Mezi tyto data patří především rozsáhlé a strukturované dokumenty jako jsou např. knihy, příručky či prezentace. Vyhledávání v takových dokumentech se provádí pomocí klíčových slov. Tyto data jsou ukládána hlavně v elektronických knihovnách.

Možné typy dokumentů jsou např. DOC, PDF, TeX, PS, HTML, XML, aj.

3.3 Metadata

Metadata popisují data (i multimediální). Digitální knihovny vyžadují přesný a strukturovaný popis každé uchovávané jednotky. Avšak digitální knihovny nejsou jediné, které tuto vlastnost potřebují. Typy metadat [12]:

- Ø **Administrativní metadata** jsou data nezbytná pro uchování, správu a následné zobrazení obsahu. Obsahují například způsob a čas pořízení, kódování, formát a velikost obsahu, nebo informace o jeho zabezpečení. Jsou záležitostí daného formátu, případně multimediálního kontejneru. Databázový systém by měl umět tyto data interpretovat a využít především pro svou interní potřebu. Pro vyhledávání jsou spíše doplňkové (celovečerní film).
- Ø **Strukturální metadata** popisují interakci objektů. Například do kterého videa patří daná stopa nebo snímek, stránka dokumentu (kompozice), jejich kategorické uspořádání (agregace), případně jejich provázání odkazy (asociace).

V této kapitole se však budu věnovat pouze popisným metadatům.

- Ø **Popisná metadata** – informace popisující zdroj, jsou viditelná uživateli a slouží především pro vyhledávání. Dle perspektivy uživatelů jsou to [13]:
 - Tvůrce obsahu vloží bibliografické informace, jako je autor, název, shrnutí.

- Poskytovatel zdroje vloží další metadata, která usnadní nalezení média v jejich kolekci. Například klíčová slova, popis obsahu, index a podobně.
- Konzument může přidat vlastní poznámky, případně požadavky (na datový tok).

Multimediální data obsahují v hlavičce souboru obvyklé informace jako rozlišení, typ, barevnost či kódování. Kromě těchto informací však mohou obsahovat ještě další informace [4]:

- Ø **ID3** je známý a používaný popis použitelný v hlavičce zvukového formátu MP3. Obsahem jsou informace jako název skladby, jméno umělce či skupiny, název alba, rok vydání nebo také žánr či skladatele.
- Ø **Exif** je standard využívaný digitálními fotoaparáty při pořizování snímků. Informace se vkládají do existujících souborových formátů jako je JPEG, TIFF a RIFF WAVE. Není však podporován ve formátech JPEG 2000 či PNG. Exif obsahuje informace jako značka a model fotoaparátu, datum a čas pořízeného snímku, nastavení fotoaparátu (citlivost, clonu, expoziční čas, ohniskovou vzdálenost atd.), náhled snímku, informace o místu pořízení (může být získána z GPS přijímače) nebo také komentáře a informace o autorovi.
- Ø **IPTC-NAA-Standard** slouží k ukládání textových informací do obrazových souborů (např. JPEG nebo TIFF) a byl definován v roce 1990 pro novinářské účely. Textové informace se ukládají v definovaném formátu do k tomu určené části souboru.
- Ø **Dublin core** je standard pro metadatový popis digitálních objektů (včetně WWW stránek), často vyjadřovaných prostřednictvím XML. Obsahuje základní sadu šestnácti prvků, z nichž žádný není povinný (název, tvůrce, předmět, popis, vydavatel, přispěvatel, datum, typ, formát, identifikátor, zdroj, jazyk, vztah, pokrytí, práva, publikum). Vzhledem k tomu, že byl vytvořen pro anotaci videam obsahuje rozšíření, které umožňuje sémanticky anotovat jednotlivé scény a podobně.
- Ø **MPEG-7**
- Ø **MPEG-21**

Toto bylo několik formátů pro popis multimediálních dat. Moderní formáty jako PNG, AAC, MPEG-4 či JPEG 2000 mají možnost vložit pouze XML metadata, tedy např. Dublin Core. Ten je však zaměřen výhradně na popis média a neobsahuje žádnou možnost vyhledávání dle obsahu. Na XML je založen i MPEG-7, který doplňuje předchozí nedostatky.

3.3.1 MPEG-7

MPEG-7 [přeloženo z 9] je standard ISO/IEC vyvinutý konzorciem MPEG (Moving Picture Experts Group), tedy tím, které stojí za úspěšnými standardy MPEG-1 (1992), MPEG-2 (1994), a MPEG-4

(1998 a 1999). Standardy kódování videa MPEG-1 a MPEG-2 jsou základem celého segmentu produktů a technologií, jako Video CD, MP3, DVD, digitální televize DVB aj.

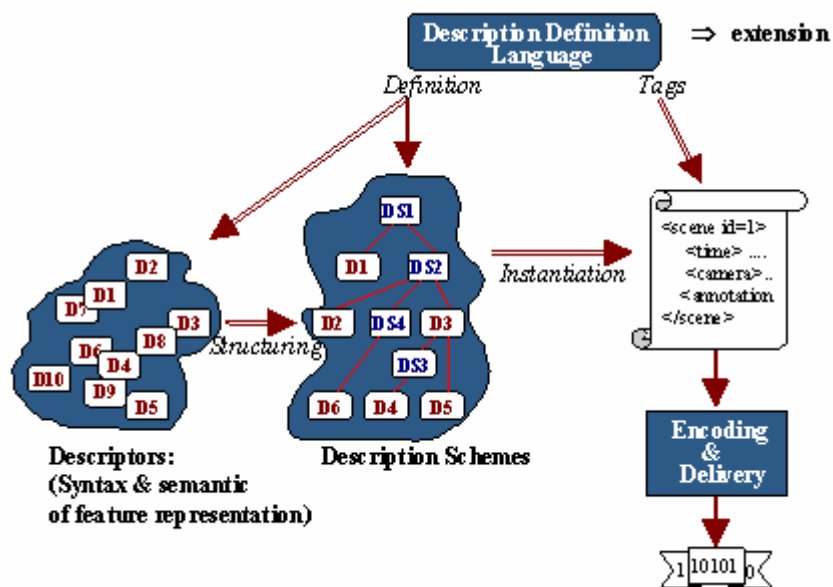
Standard MPEG-7, formálně také zvaný "Multimedia Content Description Interface", je řadou standardizovaných nástrojů pro **popis** multimediálního obsahu. V MPEG7 tedy jde o textový popis nějakého videa, např. jaký je jeho název, autor a co se v něm děje. Textový v tomto případě znamená, že je lidsky i strojově čitelný, což je dosaženo použitím značkovacího jazyka XML. MPEG-7 definuje takových nástrojů pro audiovizuální popis rozsáhlou sadu (Description Tools jsou metadatové elementy, jejich struktura a vztahy, které standard definuje ve formě deskriptorů a deskripčních schémat). Ty jsou základem pro aplikace, kterým umožní efektivní přístup k multimediálnímu obsahu.

MPEG-7 poskytuje množinu standardizovaných deskriptorů pro popis obsahu různých druhů médií – statické obrazy v tištěné podobě, 3D grafika a její modely, zvuk, řeč a video nebo biometrie lidského obličeje – s cílem efektivního vyhledávání informací ve velkém množství multimediálních dat. Popis nesouvisí se způsobem uložení médií. Je možné přiřadit deskriptor MPEG-7 k obrazu, článku, filmu stejně jako proudu MPEG-4.

Základní elementy standardu MPEG-7 jsou [9]:

- § Deskriptory (D) reprezentují vlastnosti, rysy – atributy multimediálního obsahu založené na katalogích (název, autor, práva, popis), sémantice (kdo, co, kdy a kde – informace o objektech a událostech), syntaxi (pro CBR, barva obrazu, tón zvuku) a technologii (formát, velikost, vzorkovací frekvence).
- § Popisová schémata (Description Schemes – DS) popisují strukturu a sémantiku vztahů mezi komponentami D nebo DS – typ média, jeho původ, možnosti použití, strukturální vlastnosti nebo libovolný text.
- § Jazyk pro definici deskriptorů (Description Definition Language – DDL) definuje D, DS, DT (Datové typy - zavádějí například jednoduchá pole v XML), jejich syntax, sémantiku, možnosti jejich změny a rozšíření založené na XML, upravené za pomoci W3C pro MPEG-7.
- § Systémové nástroje (Systems tools) podporují tvorbu a přenos popisů například v binární podobě, jejich multiplexování s multimediálním obsahem, synchronizaci, formáty souborů.

Na obrázku Obr. 3.1 můžeme vidět vztahy mezi elementy MPEG-7 zmiňovanými výše.



Obr. 3.1 Schéma základních elementů MPEG-7

MPEG-7 se dělí na 10 částí [9]:

1. Systém: specifikuje nástroje pro přípravu deskriptorů k přenosu, ukládání, kompresi a pro synchronizaci s obsahem.
2. DDL: specifikuje jazyk pro definování standardní množiny deskriptivních nástrojů (DS, D, DT) a nových nástrojů, založeno na XML.
3. Vizualní: obsahuje nástroje pro popis vizuální složky, specifikuje základní kategorie – barva, textura, tvar, pohyb, pozice, rozpoznávání obličejů.
4. Audio: definuje nástroje pro popis zvukové složky jako spektrální, časové, dynamické vlastnosti (low-level) nebo rozpoznání hlasu, barva nástroje, melodie (high-level).
5. Schémata popisující multimédia: nástroje pro popis multimediální části popis obsahu – audio i vizuální (video), použití obsahu, organizace, navigace, interaktivity.
6. Referenční software: implementace standardu se stále vyvíjí, protože norma popisuje způsob uložení popisu, nikoli způsob jeho získání.
7. Testování shody: poskytuje vodítka pro testování shody implementace deskriptorů.
8. Extrakce a použití MPEG-7 schémat. Pouze informativní, není součástí normy, například implementace experimentálního modelu.
9. Profily a úrovně – poskytují směrnice a profily standardu
10. Definice schématu – specifikuje schéma využívající DDL

3.3.2 MPEG-21

MPEG-21 je standard vyvinutý konzorciem MPEG (Moving Picture Experts Group) zaměřený na definování otevřeného frameworku pro multimediální aplikace. Je to nastupující standard pro komunikaci mezi všemi uživateli – tvůrci, distributory a konzumenty multimediálního obsahu v distribuovaném multimediálním systému [14]. Jednotkou dat je unikátní digitální položka (Digital Item), obsahující AV obsah (JPEG2000, MPEG-4 a další) spolu s popisem MPEG-7, správou autorských práv a systémem pro jejich distribuci, prezentaci a synchronizaci.

3.4 Ukládání

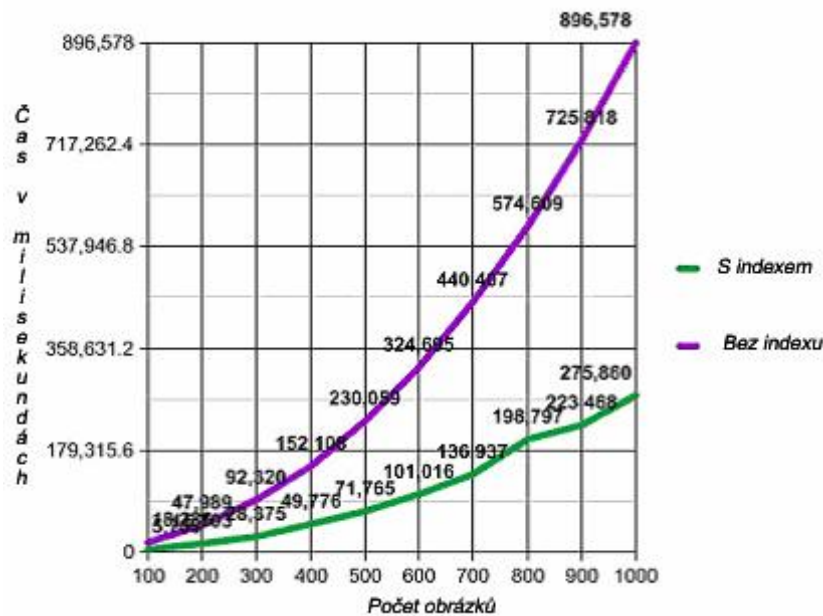
Samotné fyzické ukládání multimediálních dat klade nemalé požadavky na diskový prostor a systémové zdroje. Pro klasická strukturovaná data je umístění v datovém souboru s transakční kontrolou samozřejmé. Avšak pro multimedia je kvůli zvýšeným výkonnostním, fyzickým a finančním důvodům situace výrazně složitější. Multimediální data mohou být uloženy např. uvnitř databáze (jako objekt nebo BLOB), v externím souboru, na nějakém serveru (FTP, streamovacím, HTTP) či na nějakém zařízení jako např. kamera s dočasnou pamětí.

3.5 Indexace

Index je entita databáze, která slouží ke zrychlení vyhledávacích a dotazovacích procesů v databázi a to definováním unikátní hodnoty u všech sloupců tabulky, podle kterých se vyhledává, třídí nebo podle kterých se spojují tabulky. Existují tzv. full-textové indexy, které jsou v některých případech definovány nad celou databází. Každá tabulka může mít indexů několik a indexovat je možné jak obsah tak i popis. Způsoby uložení indexů a jejich chování se mohou výrazně lišit a to podle použité databázové technologie. Porovnání využití a nevyužití indexu při dotazování obsahu obrázkových dat lze vidět na obrázku Obr. 3.2 .

Využití indexů je často dosti zanedbáváno. U menších tabulek (pouze desítky záznamů) je význam indexů zanedbatelný avšak u větších tabulek indexy výkon dosti zásadně ovlivňují. Mohlo by se zdát, že větší počet indexů má lepší chování databáze, avšak nesmíme zapomenout na to, že každý index zabírá nezanedbatelné množství místa vyhrazené pro databázi a každý index zpomaluje operace, které mění obsah indexovaných sloupců. V případě existence indexu určuje jeho využití plánovač a to na základě optimalizovaného dotazu. Základní technikou multimediálních přizpůsobení je nezpracovávat nepotřebná data (v danou chvíli) nebo data, které nelze prezentovat.

Důležité je skutečnost, že se indexují metadata.



Obr. 3.2 Motivace využití indexu [15]

Pravděpodobně nejjednodušším způsobem indexování obsahu je využití předepsaných klíčových slov (label). V běžném textu se slova mohou opakovat a je nutné uvažovat jak synonyma, tak homonyma (řešení pomocí stromu). Nejznámější strukturou pro indexování textových dat, je frekvenční tabulka. Řádky představují jednotlivá klíčová slova (termy), což jsou všechna slova vyskytující se v uložených dokumentech (wordlist), často bez nejběžnějších slov (a, to) zvaných stoplist, nebo je použit jen kmen slova (slov). Sloupce pak tvoří jednotlivé dokumenty, respektive počet klíčových slov v daném dokumentu.

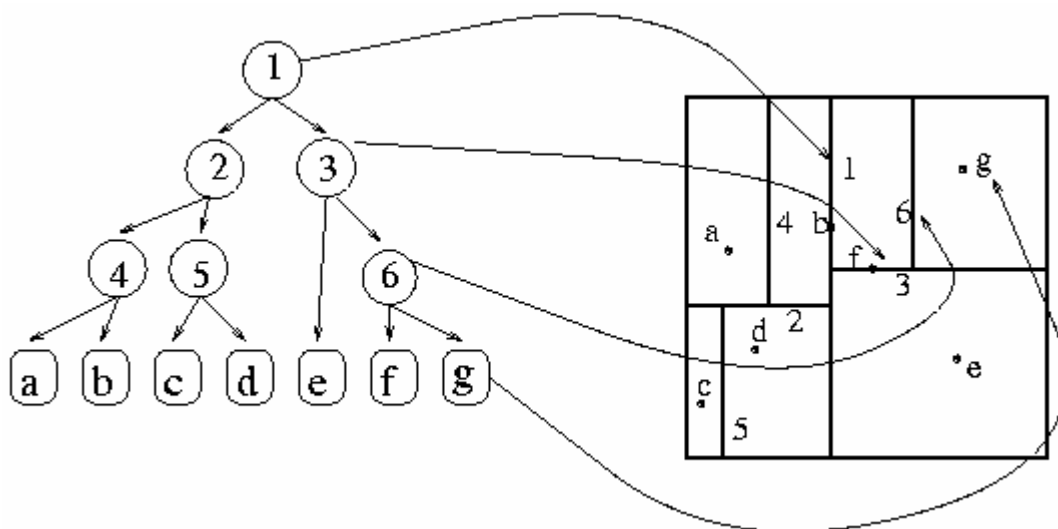
Možnost jak indexovat i další obsah je například využitím vektorů rysů. Pokud jsme schopni vytvořit vektor rysů s konstantním počtem číselných parametrů (barevný histogram, textura), je možné použít klasické indexové struktury, jako jsou B+ stromy nebo bitmapový index [10].

Tradičním požadavkem MMDBs je indexování (anotovaných) objektů. Problémem multimediálních dat je automatické nalezení těchto objektů a především jejich vícerozměrná povaha. Nelze tedy použít klasické metody, ale je možné použít struktury založené na dělení prostoru [4].

3.5.1 K-dimensionální strom (kd-strom)

Prohledávání prostoru obecně mívá vysokou výpočetní složitost. Efektivnější vyhledávání lze dosáhnout vymezením podprostoru a to např. pomocí nadrovin (rovnoběžných s příslušnými osami), vymežit část k-rozměrného prostoru, která buď obsahuje přímo hledané řešení nebo umožňuje přijatelné libovolné dohledání. K-D stromy slouží pro ukládání bodových k-dimensionálních dat.

Při stavbě stromu indexu vložíme první bod. Do levé části stromu dáme všechny další body, které mají menší hodnotu souřadnice x a do pravé větvi. Při vkládání dalšího bodu se prostor rozdělí podle další souřadnice viz Obr. 3.3. Nevýhoda K-D stromů je špatné pořadí při vkládání, výhodou je snadná implementace.



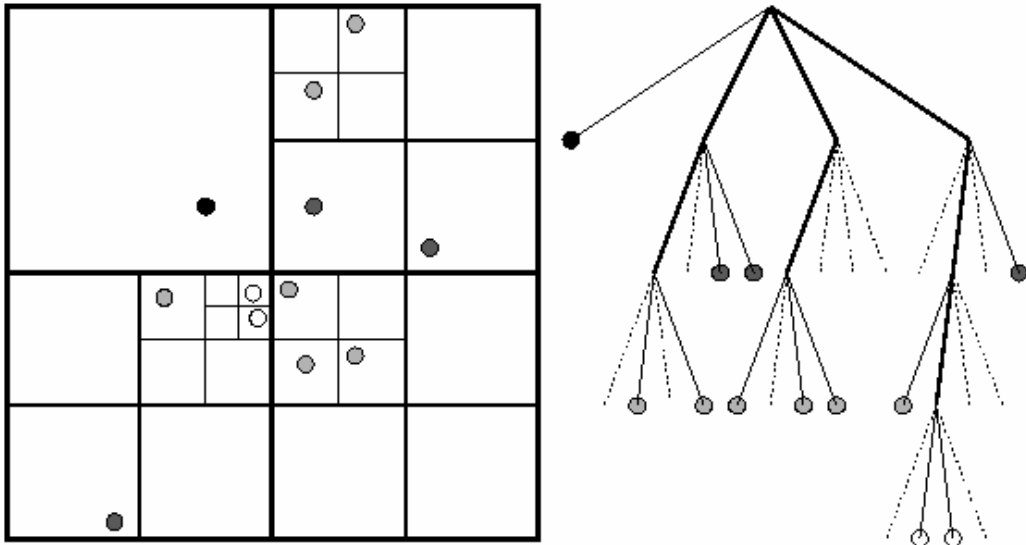
Obr. 3.3 Dělení 2D prostoru pomocí kd-stromu

3.5.2 Quad strom

Quad strom je metoda ukládání a vyhledávání záznamů z databáze. Algoritmus vyhledává data opakovaným dělením záznamů do 4 částí dokud nezůstane pouze jeden. Vychází z K-D stromu, ale vkládané body rozdělují prostor ve všech směrech, viz Obr. 3.4. Quad stromy umožňují oproti K-D rychlejší vyhledávání, ale operace rušení je náročnější, kvůli nalezení uzlu, který má být nahrazen.

Jejich modifikací jsou MX-quad stromy. Také dělí prostor podle všech souřadnic, ale jen tak, aby vzniklé části byly stejně velké. Dále dělí rekurzivně. V těchto stromech je pevně určena granularita, problémem je její efektivní určení.

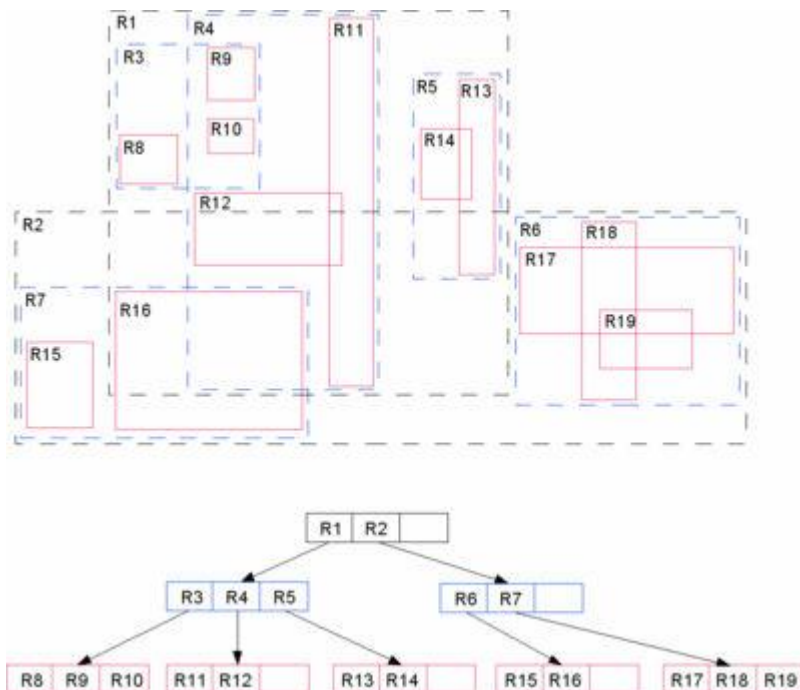
Zobecnění K-D a Quad stromu (obrázek Obr. 3.4) je BSP strom, který obdobně rozdělují prostor libovolnými rovinami na konvexní oblasti, obsahující maximálně jeden objekt, případně jeho část [4].



Obr. 3.4 Dělení 2D prostoru pomocí Quad stromu [16]

3.5.3 R-strom

R-stromy jsou datové struktury podobné B-stromům, ale využívají se např. na indexování multi-dimensionálních informací. R-stromy jsou založeny na ukládání obdélníkových struktur. Některé obdélníky představují reálné objekty vložené do databáze. Další sdružují menší obdélníky do větších, viz Obr. 3.5. Všechny uzly stromu, mimo kořenového, musí být alespoň z poloviny zaplněné. Pro neobdélníkové objekty se hledá minimální ohraničující obdélník. Modifikací R-stromů jsou M-stromy s kruhovými (kulovými) oblastmi, což je výhodnější pro určování eukleidovských vzdáleností, ale výpočetně náročnější[4].



Obr. 3.5 Dělení 2D prostoru pomocí R-stromu

R-stromy jsou nejvíce použitelné pro velké množství dat. Hlavním problémem těchto stromů je možné překrytí obdélníků. Kvůli tomu nelze při vyhledávání vyloučit některé větve, ale je nutné prohledávat větší část stromu. Vylepšenou modifikací jsou R+ stromy a R* stromy. R+ stromy vylučují překryvání obdelníků, avšak obdelník může být obsazen v několika listech. Běžnější modifikace jsou R* stromy, které při výběru uzlu pro zařazení a štěpení snaží překrytí minimalizovat [11].

Závěr

Cílem této práce bylo seznámit se s problematikou multimodálních databází a vyhledávání v nich a při prozkoumávání této oblasti jsem zjistil, že je to oblast, která je ve vývoji už spousty let a i když se od svého zniku vývoj opravdu velkým krokem posunul kupředu, tak stále není vyřešena spousta úskalí a problémů a očekávám, že vývoj ještě „několik let“ bude pokračovat.

Při řešení této práce jsem byl překvapen jak rozsáhlé a zajímavé téma to je, proto předpokládám, že při řešení diplomové práce se ještě více seznámím s danou problematikou a pokusím se navrhnout a naimplementovat takovou aplikaci nebo součást nějaké aplikace, která bude přínosem. Mohlo být např. vyhledávání informací na základě vektorového modelu.

Literatura

- [1] C. J. van Rijsbergen *Information Retrieval*. London, 1979.
- [2] Chmelař, Petr. PDB: Vyhledávání informací a multimediální aplikace [online], 2006. Dostupný z: <http://www.fit.vutbr.cz/~chmelarp/pdb/>
- [3] WIKIPEDIE : The Free Encyklopedia [online]
Dostupný z: <http://en.wikipedia.org/wiki>
- [4] Chmelař, Petr. PDB: Multimediální databáze [online], 2006.
Dostupný z: <http://www.fit.vutbr.cz/~chmelarp/pdb/>
- [5] HIGASHI, T. – KITA, S – WATANABE, I. Texture and Objects: Interruption of Same-Object Effect in Human Vision, *IEICE Trans Inf & Syst* E89-D: 1806-1812. 2006.
- [6] KOLÁŘ, D. *Postrelační databáze, Přednášky 2005*.
Dostupný z: <https://www.fit.vutbr.cz/study/courses/PRD/private/>.
- [7] Grandvalet, Y. – SANU, C. Adaptive Scaling for Feature Selection in SVMs. 2004.
Dostupný z: <http://www.hds.utc.fr/~grandval/nips02.pdf>.
- [8] KADOUS, M.W. *Dynamic Time Warping* [cit. 2006-09-10]. 2002.
Dostupný z: <http://www.cse.unsw.edu.au/~waleed/phd/html/node38.html>.
- [9] ISO/IEC JTC1/SC29/WG11. *MPEG-7 Overview* [online]. Martínez, José M. Palma de Mallorca : 2004 [cit. 2005-10-22].
Dostupný z: <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>.
- [10] ZENDULKA, J. IDS: Materiály k přednáškám ve formátu PDF [online]. 2005. Dostupný z: <https://www.fit.vutbr.cz/study/courses/IDS/private/>.
- [11] SUBRAHMANNIAN, V. S. *Multimedia database systems*. Morgan Kaufmann, 1998. 442 s. ISBN 1-55860-466-9.
- [12] Oxford Digital Library. Metadata in the Oxford Digital Library. [online]. 2005. Dostupný z: <http://www.odl.ox.ac.uk/metadata.htm>.
- [13] KOSCH, Harald. Distributed Multimedia Database Technologies Supported by MPEG-7 and MPEG-21. CRC Press, 2003. 280 s. ISBN 0-8493-1854-8.
- [14] ISO/IEC JTC1/SC29/WG11. *MPEG-21 Overview* [online]. Bormans, J. – Hill K. Shanghai: 2002.
Dostupný z: <http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm>.
- [15] Giralda, A. R. *Image Databases Indexing*. 2006. Vedoucí MT Zendulka, J.
- [16] Oxley, S. *Modelling the Capture Tudory for the Origin of Planetary Systéme*, New York, 1999