



ZZD

Získávání znalostí z multimediálních databází

Petr Chmelař

Obsah

1	Úvod	2
2	Získávání znalostí z databází z [Ha01]	3
2.1	Dolování dat	5
2.2	Multimediální databáze Oracle interMedia.....	6
2.3	Prostorové databáze	8
3	Získávání znalostí z multimediálních dat.....	9
3.1	Dolování dat z multimédií.....	10
3.1.1	Klasifikace a predikce v multimediálních datech	11
3.1.2	Asociační analýza.....	11
3.2	Způsob popisu normou MPEG-7.....	12
3.3	Aplikace	14
4	Shrnutí	17
5	Literatura	18

1 Úvod

„Proč chceme získávat znalosti z multimediálních databází?“ Říká se, že v dnešní době jsme přesyceni daty. A opravdu tomu tak je, jsme schopni až neuvěřitelně efektivně sbírat čím dál více dat. Ať už jde o ty, které potřebujeme a chceme, ale také ty, které nás nezajímají. Jde o všechny možné obory průmyslu a obchodu, také veškeré vědní obory v současné době poskytují ohromné spousty poznatků a jiných digitálních dat. Také v domácím použití lze nalézt množství dokumentů a to nemluvím o Internetu obecně.

„Topíme se v datech, ale žízníme po informacích.“

Nikdo z nás už nemá čas probírat se všemi informacemi, ke kterým máme přístup, v jejich stále se zvětšujících objemech. Zejména proto potřebujeme najít spolehlivý přístup, kterým lze data automaticky zpracovávat, klasifikovat, objevovat zajímavé vlastnosti a vztahy, změny i novinky a z nich vybírat jen to podstatné, co nás zajímá, vytvářet náhledy a výtahy.

Databázové technologie se vyvinuly z primitivních souborově-orientovaných po komplexní, objektové, dotazovací systémy. Další vývoj směřuje k efektivní analýze a pochopení uložených dat.

Tato práce se zaměřuje na multimediální data, která vytváří zvláště velké databáze, konkrétně na jejich obrazovou část, pro jejichž indexaci a vyhledávání je nutné využít některé z automatizovaných vyhledávacích technik, které pro popis obsahu obrázku využívají jejich vizuálních vlastností – rysů. Jinak se z těchto dat stávají data mrtvá, což není ani příliš přehnaná metafora, protože na druhé straně jsou rozmrzelí lidé, kteří tyto informace potřebují a chtějí, ale nemají potuchy o jejich existenci.

Na tuto výzvu odpověděla například i významná skupina MPEG, svým počinem, normou MPEG-7: Multimedia Content Description Interface [ISO1], tedy rozhraní pro popis obsahu multimedií, jehož cílem je vyhledávání v multimediálním obsahu stejně snadno jako v klasických textových datech. Také standard SQL z roku 2001 poskytuje nástroje pro nízkourovňový popis multimediálních dat na.

2 Získávání znalostí z databází z [Ha01]

„Proč chceme získávat znalosti z databází?“ Protože je potřebujeme – všechno začalo na popud obchodu a průmyslu, nejen informačního, kvůli dostupnosti ohromného množství dat a nutnosti obrátit tyto data v použitelné znalosti a smysluplné informace. Ty bylo zapotřebí použít v obchodních kruzích pro management, řízení výroby, kvality, marketingovou analýzu trhu až po inženýring a vědecký výzkum.

Je to jedna z nejméně aktivních oblastí oborů zabývajících se databázemi, do které informační technologie v poslední době dospěly přirozenou cestou. Lze zde provádět výzkum a vyvíjet nové nástroje.

Je to mladá mezioborová disciplína, která těží z databázových technologií, statistiky a matematiky obecně, strojového učení, umělé inteligence, teorií získávání a sběru dat, jejich vizualizace a efektivního computingu. Výzkum využívá poznatky také z oblastí neuronových sítí, rozpoznávání vzorů, prostorové analýzy dat, u obrazových a multimediálních databází i ze zpracování signálů; u dalších typů dat například z obchodních teorií, ekonomiky, nebo třeba i bioinformatiky a teorie informace obecně.

Získávání znalostí sestává z částí, jak ukazuje obrázek 2.1:

Čištění dat odstraňuje inkonzistentní a duplicitní data, šum i jiné anomálie.

Datová integrace kombinuje data z různých zdrojů a databází, někdy je možné považovat výsledek čištění a integrace dat jako krok předzpracování dat pro datové sklady.

Výběr dat – ze skladů jsou získávány data, která jsou relevantní pro naši analýzu.

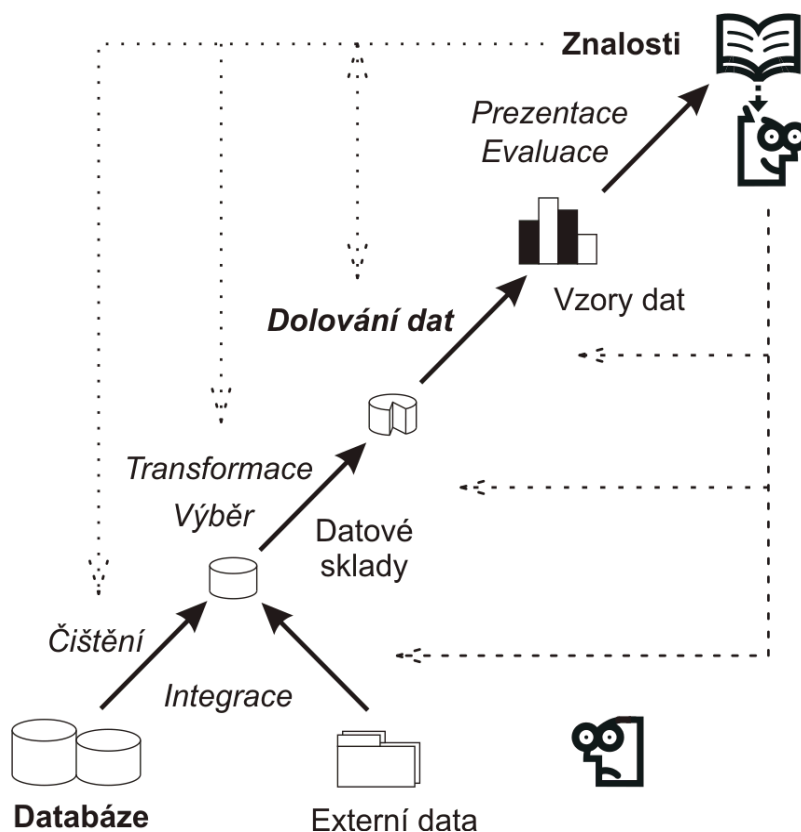
Transformace. Data jsou transformována nebo slučována do tvaru odpovídajícího pro jejich dolování. Někdy je tento krok součástí předzpracování v datových skladech.

Dolování dat je základní proces, ve kterém jsou inteligentními metodami získávány vzory dat z jejich skladů.

Evaluace vzorů slouží pro verifikaci získaných rysů, zda se jedná skutečně o požadovanou znalost, používají se pro to speciálně vyvinuté metriky.

Prezentace znalostí kombinuje vizualizační techniky s prezentačními pro zpřístupnění znalostí uživatelům.

Systémy pro získávání znalostí mohou být klasifikovány na základě typu databáze použité pro dolování, zdrojového i požadovaného typu dat, nebo podle požadovaných aplikací.



Obrázek 2.1: Proces získávání znalostí z databází

Existuje několik typů databází, podle způsobu uložení dat rozlišujeme:

Relační databáze sestává z tabulek, relací sloupců a jejich hodnot, které jsou dostupné přes systém řízení báze dat.

Transakční databáze je také soubor záznamů, ale reprezentujících transakce – jako například každý nákupní košík.

Datové sklady. Data jsou při procesu získávání znalostí obvykle uložena v datových skladech, což je depozitář, který kombinuje data z různých zdrojů, příkladem jsou sklady založené na multidimenzionální kostce OLAP. Data jsou zde uložena v normalizované podobě, obvykle bývají sumarizována, tak aby umožnila co nejefektivnější získávání informací, respektive podporu při rozhodování. Datové sklady ale mohou poskytovat i další zajímavé prostředky pro zpracování dat.

Další techniky umožňují použití objektového přístupu k výše uvedeným variantám nebo kombinovat jiné externí zdroje, jako jsou disková pole nebo Internet.

Lze také rozlišovat dle uložených dat:

Temporální a časově uspořádaná data - klasické relační a transakční databáze, převládají strukturovaná data – text a číselné informace s časovými razítky.

Prostorová data jako satelitní nebo medicínské snímky, uložené ať už jako rastr, nebo jako vektorové informace s popisem a lokalizací.

Multimediální databáze – obrazové, audio a video a jiná heterogenní data.

Textové databáze dokumentů jako XML/HTML s dalšími, potenciálně různorodými informacemi až po WWW.

Hlavní problémy při dolování dat řešené v současné době a body vhodné pro možný další výzkum, vývoj a jeho zlepšování:

Metodologie dolování a interakce s uživateli.

Protože různí uživatelé požadují různé druhy znalostí, je nanejvýš vhodné ze stejné databáze dolovat různými metodami s různými parametry na různých úrovních abstrakce, dolování by mělo být co možná nejvíce interaktivní, protože někdy může být těžké předem odhadnout co všechno lze získat.

Efektivní by dále bylo vytvoření univerzálního dotazovacího jazyka pro získání znalostí a ad hoc dotazy, který bude moci obecně poskytovat odpovědi, dedukce a případně i jiné znalosti na mnoha úrovních.

Také prezentace a vizualizace výsledků musí být použitelná a snadno pochopitelná uživateli. Manipulace se zašumělými a neúplnými daty může také způsobit zmatení metody a nepřesné výsledky vyhledávání.

Ohodnocení nalezených vzorů – problém zajímavosti, systém je schopen nalézt miliony vzorků, přičemž většina z nich je pro daného uživatele naprosto nezajímavá.

Problémy týkající se efektivity a celkového výkonu aplikace vzhledem k obrovským objemům zpracovávaných i generovaných dat.

Body týkající se diverzity databázových objektů. Velký prostor pro zkoumání a dolování si zaslouží databáze, které obsahují více komplexního obsahu, například multimediálního. Způsob dolování dat, postavený na jednom typu logicky nebude fungovat pro různé typy objektů. Je nutné vytvořit nové, speciální metody pro specifické aplikace.

Posledním problémem je bezpečnost dat a sociální dopad výsledků.

Výsledkem je, že nejsme schopni dobrat se opravdu kvalitních výsledků dříve než v řádu let. Musíme objevit přístupy, které jsou dostatečně efektivní, snadno realizovatelné, použitelné a plně přizpůsobitelné různým oblastem zájmu pro vyhledání pouze plně relevantních dat z jejich skladů.

2.1 Dolování dat

Výraz dolování dat je trochu zavádějící, pro lepší pochopení je to vlastně naopak: dolování v datech. Je zde míněn proces ve smyslu objevování zajímavých vzorů z velkých objemů dat uložených v databázích, datových skladech, nebo jiných skladech informací, jako jsou data každého uživatele nebo Internet.

Je to esenciální krok při získávání znalostí z databází, i když někdo může oba pojmy zaměňovat. Obrázek 2.1 znázorňuje jejich vztah.

Z perspektivy datových skladů je dolování dat jen jako rozšířená fáze analytického zpracování OLAP. Ve skutečnosti je ale dolování dat mnohem obecnější než analýza agregačně-sumarizačního stylu OLAP.

Znalosti prezentované vzory jsou pro lidi poměrně snadno pochopitelné, musí být platné na analyzované množině dat s jistou pravděpodobností, potenciálně použitelné, nové, nebo skrývat to co je nezajímavé nebo raritní, pokud si to uživatel nepřeje naopak zdůraznit.

Lze rozlišit dva hlavní směry při dolování dat:

Popisné dolování objevuje zajímavé vlastnosti dat.

Prediktivní dolování popisuje současná data s důrazem na hypotézy vývoje vlastností v budoucnosti.

Existuje několik typů úloh, kterými lze tyto cíle dosáhnout v závislosti na typu dat a požadovaných vzorků:

Diskriminace a charakterizace rozřazuje data podle jejich příslušnosti do tříd, nebo vytváří vazby mezi jejich skupinami. Například myš a tiskárna patří do skupiny periferie. Charakterizace je sumarizace zajímavých vlastností dané třídy dat, například chování zákazníků s největším odběrem zboží. Bývá součástí OLAP.

Asociační analýza se snaží nalézt vztahy mezi analyzovanými daty. Ty se vyjadřují asociačními pravidly tvaru implikace $A \Rightarrow B$. Společně s metrikami podpora $S(A \Rightarrow B) = P(A \cup B)$ a spolehlivost $C(A \Rightarrow B) = P(B | A)$ určují které hodnoty atributů se vyskytují v dané množině dat s jistou pravděpodobností společně. Více v [C205, Ha01].

Klasifikace a predikce je proces, který hledá modely, popisující a odlišující známé třídy dat, tak aby bylo možno zařadit i neznámé objekty, respektive určit jejich trendy. Používají se zde rozhodovací stromy, matematická logika, nebo neuronové sítě pro sledování změn vzorců chování objektů v čase.

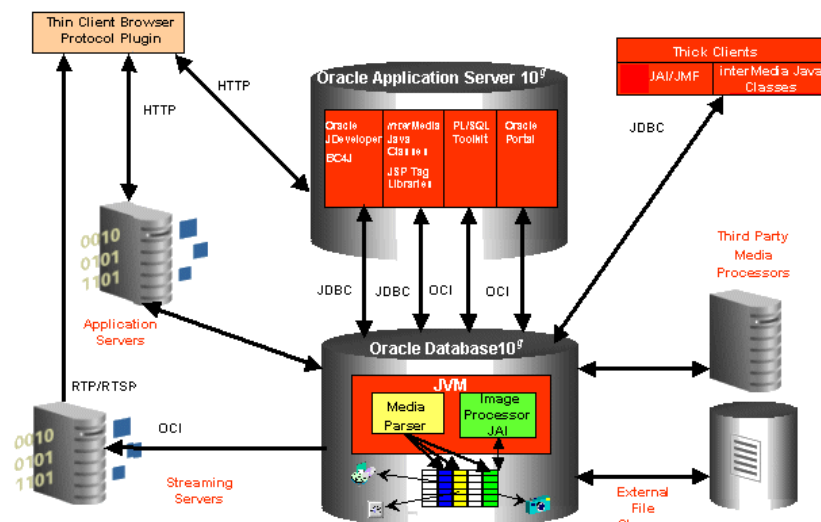
Shlukování nebo segmentace se na rozdíl od klasifikace snaží data zařadit do skupin, které předem nezná. Objekty jsou shlukovány do tříd na základě podobnosti, respektive odlišnosti mezi jednotlivými jejich vlastnostmi. Například obrázek 2.2, nebo lze třeba segmentovat zákazníky dle jejich průměrné hodnoty měsíčního odběru u dané společnosti.

Okrajová analýza naopak snaží hledat nezvyklé vzory v chování dat, které jiné metody označí jako šum. Používají se metriky pro měření odchylky od tříd, stanovených metodami výše, třeba pro odhalení nekalého záměr nebo činu. V prostoru viz. obrázek 2.2.

Evoluční analýza popisuje „zvyky“ objektů, jejich změnu v průběhu času a snaží se odhalit jejich trendy. Může zahrnovat všechny výše zmíněné postupy tak, aby se například dokázala mapovat na ceny akcií.

2.2 Multimediální databáze Oracle interMedia

Oracle interMedia je rozšíření databázového systému pro uchování, získávání, správu a dotazování obrázků, audia, videa, podobných heterogenních a jim přidružených dat pro libovolné aplikace.



Obrázek 2.2: Schéma Oracle interMedia

Oracle interMedia je založeno na objektově relačním SŘDB. Obsahuje:

Zdroj a jeho umístění (i mimo databázi) MM ve všech možných formátech.

Formát MM a jiných, heterogenních, dat. Pro statické obrázky rozměry, rozlišení, počet barev, kompresní formát. U zvukových nahrávek použité kódování, počet kanálů, vzorkovací frekvence a doba trvání. Video má rámce popsané jako obrázek, jejich počet za sekundu, kompresní algoritmus, celkové trvání a datová propustnost, případně typ MIME.

Metadata získaná z externích zdrojů, například autor, copyright, což jsou informace extrahované z jiných XML metadat a zpracovává zabudované informace, které jsou součástí MM jako je EXIF, IPTC-IIM, XMP nebo medicínské DICOM.

Oracle ve verzi Enterprise obsahuje navíc **nástroj pro popis statických obrázků** dle jejich obsahu. Pro jejich indexaci a podobnostní vyhledávání. Vektor rysů (signature) obsahuje následující nízkourovňové elementy, které extrahuje při ukládání do DB:

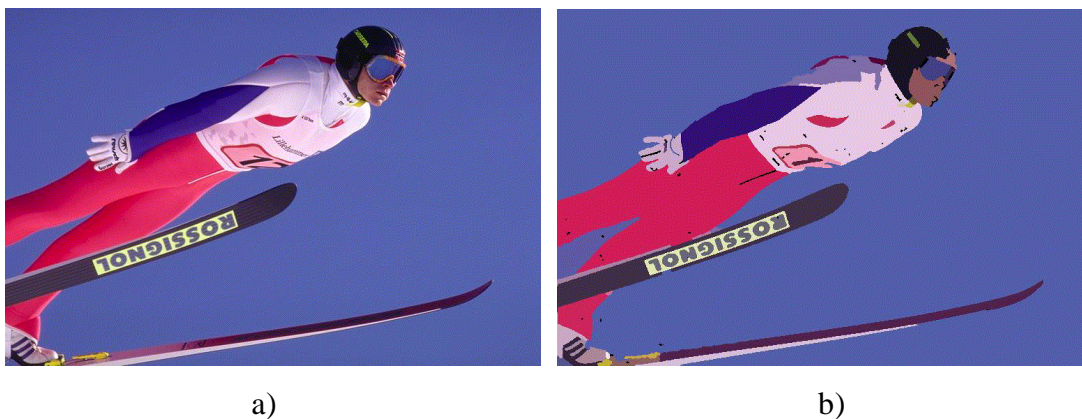
Barva, respektive její distribuce v obrázku – histogram.

Textura charakterizuje zrnitost obrázku a hranovou frekvenci.

Tvar objektů, získaných na základě segmentace obrazu, provedené podle barvy. Pro každou oblast jsou dále definovány texturní vlastnosti a ...

Umístění barevných segmentů v obraze. Pomocí této vlastnosti se implicitně nevyhledává, ale slouží například k určení pozadí, jak ukazuje obrázek 2.3.

Tento popis slouží pro podobnostní vyhledávání, realizované podle vážené vzdálenosti jednotlivých elementů v databázi, jejichž celková velikost je 3-4kB.



Obrázek 2.3: Znáznorňuje segmentaci do oblastí podle jejich barvy, z [Or05]

InterMedia je přístupné pomocí jazyků Java a PL/SQL, příklady jejich využití jsou uvedeny v [Or05]. Nabízí také jednoduché prostředky pro manipulaci s MM obsahem, například konverzi formátů, ořezání, změnu rozlišení, nebo gama korekce. Možné je rozšíření o další speciální algoritmy pro zpracování multimedií, jako jsou kodeky nebo subsystém pro rozpoznání řeči.

2.3 Prostorové databáze

Prostorové databáze obsahují množství dat s nějakým vztahem k prostoru, jako jsou mapy, zpracované snímky vzdáleného pozorování, medicínská obrazová data nebo návrhy mikročipů.

Na rozdíl od klasických relačních databází mají spolu s multimediálními mnoho společných odlišností. Obsahují klasické objekty (záznamy) včetně jejich topologické struktury – tvar oblastí, jejich sousedství a na druhé straně informace o vzdálenostech a rozměrech daných 1 – 3 rozměrných objektech – body, přímky, plochy (polygony) a objemy, které jsou uloženy v sofistikované struktuře prostorových databází. Pro jejich dotazování se používají jazyky s prostorovým rozšířením, obvykle ale zachovávají kompatibilitu s SQL.

Dolování prostorových dat se provádí za účelem získání informací o tvaru, sousedství nebo vzdálenostech a mírách, které nejsou explicitně uloženy v databázi. Dolování se snaží o lepší pochopení prostorových dat, umožňuje nalézt vztahy mezi prostorovými a neprostorovými informacemi uloženými v klasických databázích, jejichž vztahy by se bez této informace nedaly nalézt.

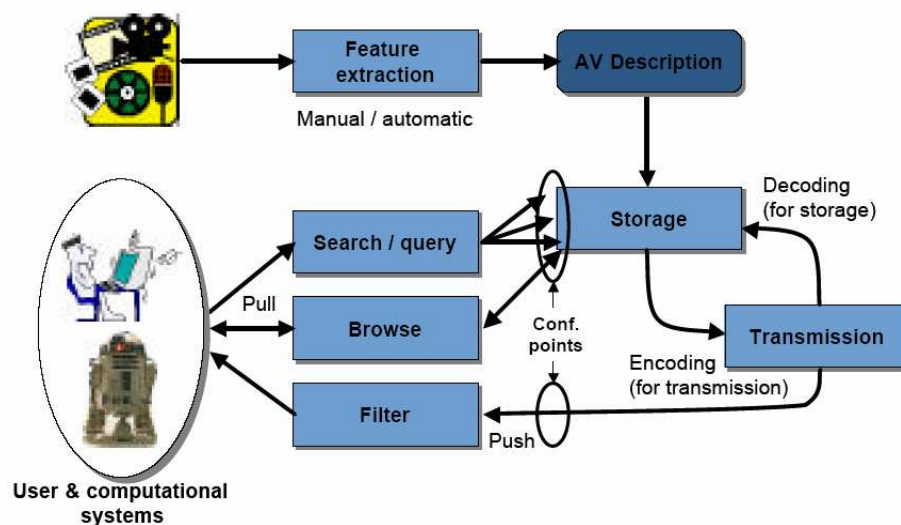
Například prostorová analýza trendů, umožňuje sledovat změny dat v prostoru – jak se mění osídlení vzhledem ke vzdálenosti od moře, nebo v čase – například lokální předpovědi počasí, změna klimatu nebo vegetace v čase a vzdálenosti od moře.

Námětem pro další práci v této oblasti je vytvoření vhodného rámce pro získávání, uchovávání a vyhledávání v časoprostorových datech (spacio-temporal databases) umožňující následně efektivní získávání znalostí z multimedií, které by byly zdrojem těchto časoprostorových dat.

Například sledování pohybu aut ve městě je možné vylepšení dopravního značení, pomůže tak k umístění parkovišť, zvýšení propustnosti a tím i snížení nehodovosti.

3 Získávání znalostí z multimediálních dat

Multimediální databáze obsahují a spravují velké kolekce multimediálních objektů, jako je audio, statický obraz, jeho sekvence – video nebo například hypertext. Význam multimediálních databází roste s popularitou zařízení pro jeho prohlížení – například digitální foto, CD/DVD video nebo rychlý Internet. Typickým příkladem takovéto databáze je systém NASA EOS (Earth Observation System), který obsahuje obrazové, ale i audio-video záznamy, databáze lidského genomu apod. Schéma je uvedeno v obrázku 3.1.



Obrázek 3.1: Neformální schéma, převzato z [M704]

Z pragmatických důvodů, popsaných výše, je nanejvýše vhodné multimediální data popsat tak, aby bylo umožněno jejich snadné vyhledání. Nabízejí se dva přístupy:

Dle popisu dat (description-based retrieval). Jedná se o deskripci sémantiky dat, která je nutné nějakým způsobem popsat – vytvořit název, klíčová slova, pojmenování osob, objektů, obrazů a akcí – to nelze obecně provést žádným automatizovaným systémem. Popis musíme vytvořit sami při jejich vzniku a nebo například při systematickém procházení a třídění, což není příliš jednoduchý, rychlý ani levný způsob, nehledě na neúplnost a inkonzistenci způsobenou subjektivním přístupem.

Automatický popis obsahu (content-based retrieval). Jedná se o metodu popisu syntaxe, plně nebo alespoň z velké části automatizovatelnou. To bohatě nahrazuje její nedostatky, zejména co se týká kvality popisu sémantiky. V případě zvukových dat se může jednat o rozpoznání klíčových slov, nástrojů, not, tempa nebo i žánru. Při popisu obrazových dat, máme k dispozici histogram barev, hrany, tvary, objekty, textury, umístění, pohyb, popis obličeje nebo jiné biometrické informace.

Jejich kombinace. Je používána zejména, pokud známe název snímku, datum pořízení, případně nějaká klíčová slova a z jeho obsahu mohou být extrahovány další vlastnosti.

Popis na základě obsahu využívá vlastností vizuálního vnímání lidí, zejména nalezení podobnosti, což je v mnoha aplikacích velice žádoucí. Ve vyhledávání zaměřeném na obsah se obvykle používají dva typy dotazů:

Předložení vzorového obrazu. Využívá se při tom porovnání deskriptorů (např. vektorů, popisujících vlastnosti) extrahovaných z obrázku dotazu s deskriptory obrazů, které byly předem extrahovány a uloženy do databáze. Výsledek dotazu je několik obrázků z databáze, jejichž deskriptory jsou si nejbližší, a tedy by měly být tomu dotazovanému nejpodobnější.

Specifikace vlastností (query-by-example). Můžeme vytvořit náčrtek, zadat barvu, tvar, specifikovat texturu, nebo zanotovat melodii, ze které je opět extrahován deskriptor a je porovnán s těmi, které jsou uloženy v databázi.

Vyhledávání zaměřené na obsah má v reálném světě mnoho aplikací, jako například pomoc při diagnóze v medicíně, předpověď počasí, vyhledávače obrázků na Internetu a samozřejmě také v libovolném průmyslovém či vědním odvětví.

Proces získávání znalostí je obdobný jako je znázorněno na obrázku 2.1, ale jednotlivé kroky je nutné doplnit o krok předzpracování – čištění (odstranění šumu), transformace multimediálních dat a extrakce jednoduchých rysů. Provádí se zejména aplikace různých filtrů pro získání vizuálních rysů, následně jejich shlukování a detekce případné změny v prostoru obrazu a času videa. Takovýto deskriptor je uložen do datových skladů pro jednoduché vyhledávání a další zpracování.

Získávání znalostí z multimediálních lze také rozdělit dle rovin aplikace:

Obecné aplikace se snaží získávat znalosti z univerzálních vstupních dat, například www obrázků, o kterých nemáme žádnou apriorní znalost a nejsme schopni odvodit mnoho o sémantice. Například by znalostní báze musela pokrýt vzory všech známých objektů v libovolné formě, jak je uvedeno v [Ar30].

Speciální využívá účelových zdrojových dat – můžeme říci, kde byly získány, a je možné stanovit, jaké znalosti lze dolovat zvoleným přístupem, protože sémantika je alespoň částečně známa předem.

3.1 Dolování dat z multimédií

Problém je, že obecná data jsou příliš rozmanitá. Objekty, tak jak je vnímá člověk, jsou velice odlišné od způsobu jejich vnímání počítači, tedy oblastí s jistou texturou, hranami, tvary, jejich přerušení, jistou trajektorií pohybu a podobně, mohou mít různou velikost i orientaci, částečně také barvu a přitom mít stejný význam, např. auto, nebo právě naopak.

V tom je kámen úrazu – například v kapitole 9.3.2 [Ha01] je zmíněno dolování typu soboru (jpeg/gif/mpeg), rozlišení (velikost) obrazu/video, obrazu, jeho framerate a podobně, což není příliš vhodné pro získávání znalostí, protože to s nimi jednoduše nemá nic společného.

Existuje několik jednoduchých přístupů k extrakci rysů z multimediálního obsahu. Zaměřím se zejména na obrazovou informaci:

Dle barvy a jejího histogramu – obsahem deskriptoru je histogram (kompozice) barev, který neobsahuje žádné informace o objektech obrázku, jako jejich rozmístění, tvar nebo textura, proto mohou být obrázky s podobným barevným rozložením naprosto odlišné.

Složený deskriptor je tvořen součástmi jako je barevný histogram, rozložení objektů, jejich tvar a textura. Ani tato kombinace obvykle nedává úplné výsledky, kvůli nesourodosti a vzdálenosti lidskému vizuálnímu systému, nicméně může sloužit jako vhodné předzpracování dat. Vyhledávání je tvořeno nalezením vzdáleností mezi jednotlivými vektory rysů.

Vlnková transformaci a přístupy na ní založené, využívá (dominantních) koeficientů vlnkové, fourierovy nebo kosinové transformace, např. v [Ch04]. Vlnky zachycují tvar, texturu i lokaci informace v obrázku.

Regionálně vymezená vlnková transformace je technika zajišťující výpočet granularity oblastí. To je založeno na poznatku, že obrázky, které by člověk označil za podobné, obsahují pouze podobné objekty, nezávisle na jejich umístění, měřítku nebo otočení. Dotaz probíhá tak, že se provádí porovnání významných částí z obrázku s fragmenty předpřipravenými v databázi a snaží se najít jejich páry nehledě na jejich umístění, velikost či orientaci.

Pohyb a jeho klasifikace se také začínají objevovat v tomto kontextu, protože dokáže odlišit pohybující se objekty od pozadí. Problémem je náročnost jeho získání a také rozlišení typu, může se jednat například o pohyb kamery.

Bylo vyvinuto velké množství technik založených na modelování, geometrii či topologii obrazové informace, nicméně jejich výčet není předmětem této práce.

3.1.1 Klasifikace a predikce v multimediálních datech

Klasifikace multimediálních dat využívá odnože umělé inteligence a počítačového vidění pro úzký okruh speciálních, poměrně efektivních aplikací.

Příkladem jsou astronomické systémy, které mají k dispozici modely známých hvězd, galaxií a jiných vesmírných objektů. Na druhé straně obrazová data z teleskopů, která dle oblasti výskytu, magnitudy, intenzity, momentu a orientace analyzují a přiřazují jim známé objekty, s tím, že hlásí ty, které známé nejsou nebo mají nějak zvláštní vlastnosti či chování. Více v kapitole 7 [Ha01].

3.1.2 Asociační analýza

Asociační pravidla mohou být získávána i z databází audia, ale zaměřím se na statické a pohyblivé obrázky:

Vztahy mezi obsahem a popisem multimédia jsou pravidla typu: „Pokud je značná část horní poloviny modrá, zřejmě se jedná o oblohu.“ Tato kategorie je sblížením obsahově a popisově zaměřeného získávání informací, protože

z obsahu obrázku nebo scény usuzuje na klíčové slovo a tím je obloha. Nebo naopak při klíčovém slově „ne“ lze z videa vyzorovat mírný pohyb hlavou.

Asociace obsahu dat bez určení jejich umístění – do této kategorie náleží například výroky jako: „Pokud obrázek obsahuje dva modré čtverce, pak je zde obvykle i žlutý kruh.“ Protože se jedná o popis obsahu obrázku, ale žádný objekt zde není lokalizován.

Asociace obsahu dat se specifikací jejich lokalizace: „Pokud je červený trojúhelník mezi dvěma modrými čtverci, pak jsou všechny umístěny ve žlutém kruhu,“ náleží k této kategorii, protože objekty jsou asociovány pomocí jejich prostorových vztahů.

Při dolování asociačních pravidel z multimedií můžeme každý multimediální objekt, například obrázek, považovat za transakci pro nalezení frekventovaně se nacházejících vzorů dat, nahrazujících položky.

Oproti klasickému modelu asociačních pravidel je zde ale několik rozdílů – každý obrázek obsahuje hodně objektů, které mohou mít mnoho vlastností jako je barva, tvar, textura, umístění, resp. obsahovat svoje klíčová slova – to znamená velmi mnoho možností pro vytvoření možné asociace. V mnoha případech je možné tutéž vlastnost ve dvou obrázcích nalézt ve značně odlišné formě, například relativní velikosti – je nutné uvažovat o tzv. progresivním upravení velikosti, protože v závislosti na velikosti je obvykle změna důležitosti objektů, oblastí na obrázcích, jako je obličej na průkazové fotografii oproti záběru na rušnou ulici.

To naráží také na problém míry opakování daného objektu. Jedná se o něco odlišného oproti klasickému konceptu asociačních pravidel použitých při analýze nákupního košíku se dvěma mléky v transakci, jejichž počet lze bez problémů zanedbat, kdežto počet pixelů v obraze ne.

Posledním problémem je zachycení uspořádání obrazu – definují se prostorové pojmy jako u, nad, vlevo, mezi, za, ... pro vztahy mezi objekty. Ty se přidávají k základním vlastnostem jako barva, tvar, textura a tvoří topologickou strukturu v prostoru.

3.2 Způsob popisu normou MPEG-7

ISO/IEC 15938: Multimedia Content Description Interface je rozhraní pro popis obsahu multimedií. Bylo sice vytvořeno skupinou MPEG, Motion Pictures Expert Group, tedy skupinou expertů na pohyblivé obrázky, ale neslouží pro kompresi audiovizuálních dat.

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 lidské obličej – s cílem efektivního vyhledávání informací ve velkém množství multimediálních dat.

Popis nesouvisí se způsobem uložení médií. Je možné přiřadit deskriptor MPEG-7 ke klasickému filmu, článku v časopise, stejně jako proudu MPEG-4.

Pro uložení schémat deskriptorů je definován formát založený na XML, který lze převést do efektivní binární podoby. Norma specifikuje prvky:

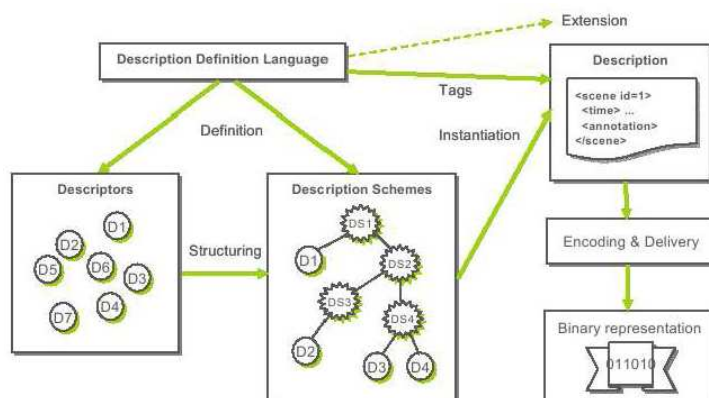
Deskriptory (D) reprezentují vlastnosti, rysy – atributy multimediálního obsahu založené na katalozích (název, autor, práva), sémantice (kdo, co, kdy a kde – informace o objektech a událostech) a struktuře (barva obrazu, tón zvuku).

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.

Datové typy (DT). Např. zavádějí jednoduchá pole do XML.

Jazyk pro definici deskriptorů (Description Definition Language – DDL) definuje D, DS, DT, 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...



Obrázek 3.2: Schéma prvků MPEG-7, převzato z [1]

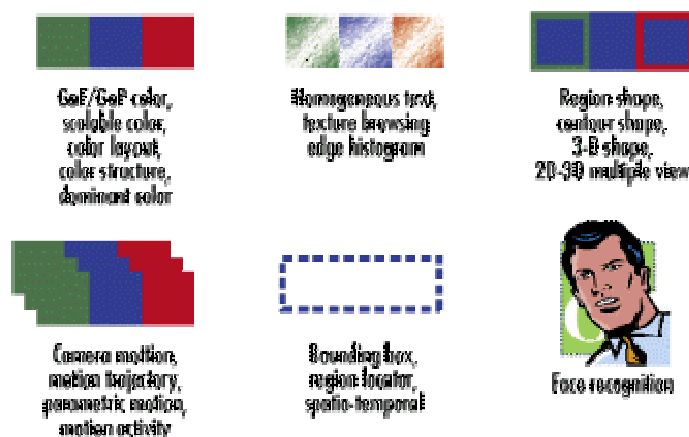
Norma MPEG-7 se dělí se na 7 tématických částí:

Část 1 – Systém: specifikuje nástroje pro přípravu deskriptorů k přenosu, ukládání, kompresi a pro synchronizaci s obsahem.

Část 2 – DDL: specifikuje jazyk pro definování standardní množiny deskripčních nástrojů (DS, D, DT) a nových nástrojů, založeno na XML.

Část 3 – Vizuální: obsahuje nástroje pro popis vizuální složky, specifikuje základní kategorie – barva, textura, tvar, pohyb, pozice, rozpoznávání obličeje.

¹ Pokud by na internetu fungovalo vyhledávání podle obsahu, byla by hračka zjistit, z jakého zdroje pochází tento obrázek...



Obrázek 3.3: Extrakce různých vizuálních rysů, převzato z [M702]

- Část 4** – Audio: definuje nástroje pro popis zvukové složky jako spektrální, časové, dynamické vlastnosti (low-level) nebo rozpoznání hlasu, barva nástroje, melodie (high-level).
- Část 5** – Schémata popisující multimedia: nástroje pro popis multimediální části popis obsahu – audio i vizuální (video), použití obsahu, organizace, navigace, interaktivity.
- Část 6** – Referenční software: implementace standardu je ve vývoji. Norma popisuje způsob uložení popisu, nikoli způsob jeho získání, na čemž se usilovně pracuje, mnohé už je ale vytvořeno.
- Část 7** – Testování shody: poskytuje vodítka pro testování shody implementace deskriptorů.
- Část 8** – Extrakce a použití MPEG-7 schémat. Pouze informativní, není součástí normy, například implementace experimentálního modelu.

3.3 Aplikace

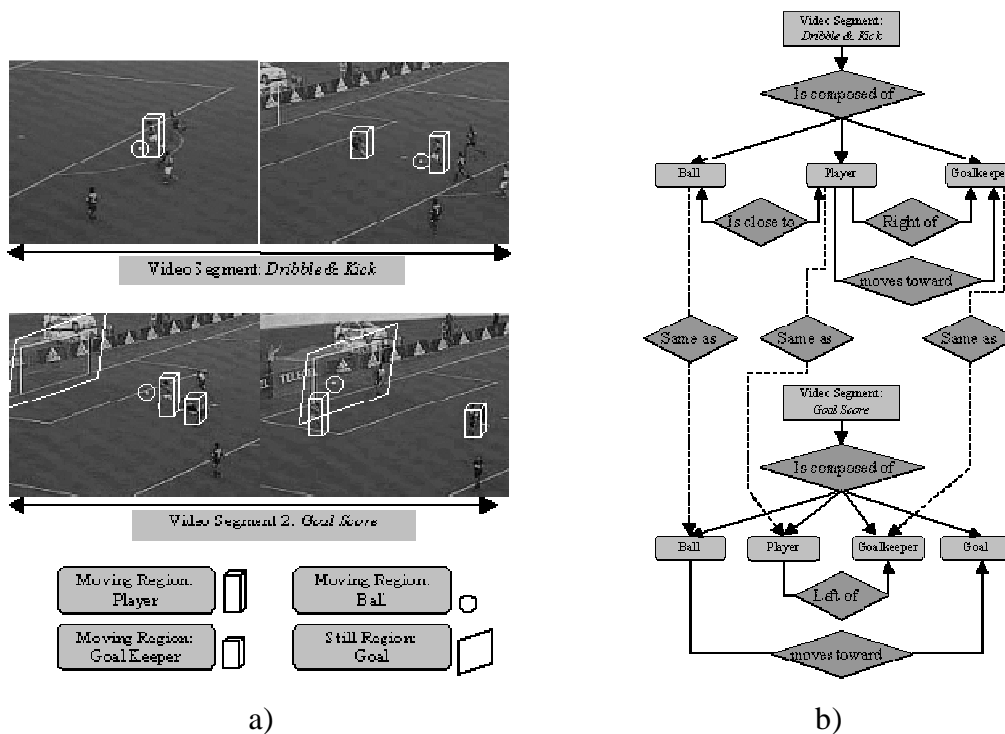
Norma MPEG-7 specifikuje pouze to, do jaké podoby se má zakódovat daná vlastnost, nikoli způsob, kterým lze tyto výsledky získat. Nicméně předpokládá 2 úrovně popisu:

Nižší (low-level) je navržena pro automatizované získávání strukturních rysů, jako je barva, pohybující se oblasti nebo strukturální vlastnosti zvuku. Uživatelé mohou například zadat logo společnosti a vyhledat všechny jejich materiály.

Vyšší úroveň (high nebo deep-level) již obsahuje volnější deskriptory pro sémantiku, kterou jsou obecně schopni specifikovat pouze lidé. Jinak je tomu ve speciálním případě, jak ukazují například obrázek 3.3.

Například získávání znalostí z videozáznamu fotbalového utkání je poměrně speciální záležitost, do jisté míry je možné popsat fotbalový zápas jednoduchými pravidly, kterým je schopen porozumět jak fanoušek Baníku, tak i počítač. Kamery zabírají reálnou zelenou plochu, která je rozdělena a ohraničena bílými liniemi. Mimo ně nás záznam obvykle nezajímá. Na hřišti je několik typů objektů – hráči

(2x10 odlišné barvy), statické branky s brankáři libovolné (šedé) barvy a míč se specifickými vlastnostmi. Rozhodčí, případně ostatní postavy klasifikovat nebudeme. Ústřední událostí je míč v brance. Předmětem získávání znalostí z multimediálních a obecně časoprostorových databází je například to, kdo se snaží dostat míč kam, jakým způsobem se mu to daří, kdo přihrává komu, jak dlouho drží míč, kde a vyvodit z toho jisté závěry.



Obrázek 3.3: Příklad a) rozpoznání objektů videa a b) jeho reprezentace z [M704]

„Jak lze dosáhnout popisu takového popisu z obrázku 3.3?“ Extrakci těchto rysů lze rozložit do několika částí:

Definice zdrojů a cílů, je nutné vytvořit syntaxi (jaké objekty se v obraze mohou vyskytovat, jak vypadají, vytvořit jeho modely, ...) a navázat na ně sémantiku (co daný objekt reprezentuje) informací, které se v obraze mohou vyskytovat. Příkladem je popis o odstavce výše.

Extrakce objektů – každý pohyblivý region může být označen za objekt a/nebo rozdělen do podmnožin, které budou analyzovány samostatně, ať už se týká vektoru pohybu (prolínání hráčů, ruce, nohy) nebo jejich popisu jako statických obrázků. MPEG-7 poté nabízí rámec pro uložení těchto časoprostorových dat a jejich vzájemné závislosti.

Transformace do souřadnic reálného světa – definice a vytvoření snímaného prostoru, specifikace polohy a pozice senzorů (kamer) a transformace detekovaných objektů v obraze do reálných souřadnic za pomoci stacionárních bodů na hřišti (lajny, branky).

Vytvoření databáze pohybujících se objektů na různých úrovních – dle MPEG-7, ale je možné vytvořit i časoprostorovou databázi a do ní ukládat pohyby

jednotlivých objektů tak, aby z nich bylo možné snadno vyvodit další informace.

Z této databáze lze vyvozovat znalosti následujícími přístupy:

Diskriminace a charakterizace rozřazuje hráče podle jejich příslušnosti do týmů.

Asociační analýza dokáže například odhadnout, že pokud při zápasu Sparta – Arsenal nahraje míč Pirés (7) Henrymu (14), pak padne do soupeřovy branky s pravděpodobností 66%.

Klasifikace a predikce. Může sloužit na jedné úrovni pro estimaci pohybu a jeho vyhodnocení například v okamžiku, kdy kamera nezabírá dané objekty, nebo na globální úrovni může odhadnout výsledek zápasu Brno – Liberec na 0 : 2.

Shlukování má obecně široké uplatnění, může vytvořit zajímavé srovnání osobní obrany a vytvořit nové skupiny hráčů například při standardních situacích.

Okrajová analýza je zřejmě schopná rozpoznat různé nestandardní situace, jako je zranění hráčů nebo změna taktiky.

Evoluční analýza naopak popisuje obvyklé chování objektů a mohla by například být schopná odhadnout komu nahraje daný hráč a s jakou pravděpodobností nebo odhadnout výsledek utkání podle první poloviny.

4 Shrnutí

Tato práce nabízí úvod do získávání znalostí z databází, jejich stručnou klasifikaci dle spravovaných dat a dále se zaměřuje na multimediální databáze, zastoupené například Oracle interMedia, s odbočkou k prostorovým databázím.

Vytvoření popisu pro získávání znalostí z multimediálních dat je možné manuálně nebo automaticky, podle obsahu. Druhý způsob se zdá být efektivnější vzhledem k množství dostupných dat. Je zmíněno dotazování na základě předložení hledaného objektu nebo specifikací vlastností pomocí příkladu. Zaměření této práce je zejména na vizuální informace.

Na úvod technik pro dolování v multimediálních datech jsou zde uvedeny některé možnosti, pomocí kterých je nutné data popsat aby z nich poté bylo možné pomocí klasifikace a predikce, nebo asociační analýzy získat další zajímavé informace.

Je zde také naznačeno, že je možné použít obecné techniky pouze do jisté úrovně, protože takové algoritmy nejsou ve většině případů schopné rozpoznat všechny objekty, tak jak jsou vnímány lidmi. Z tohoto důvodu jsou zmíněny speciální aplikace. Ty mají předem definovanou množinu objektů, které se snaží identifikovat a je možné z nich vyvozovat mnohem zajímavější informace, než z obecných, což demonstruje příklad na závěr této práce.

Toto rozlišení je zohledněno i normou MPEG-7. Ta definuje vlastnosti, které je možné popsat na nižší (barva, textura, tvar, pozice, pohyb) i vyšší úrovni (globální pozice, pohyb, popis objektů a jejich sémantika). Norma také specifikuje způsob transformace těchto vlastností do formátu založeného na XML pro snadné vyhledávání a získávání nových, zajímavých a potenciálně užitečných znalostí.

5 Literatura

- [Ar30] ARISTOTELÉS ze Stageiry. *Metafyzika*. Athény: kolem 300 př.n.l.
- [C205] CHMELAŘ, Petr. *Získávání znalostí z databází a jeho aplikace v texturní analýze*. FIT VUT Brno, 2005. Vedoucí diplomové práce Martin Heckel.
- [Ha01] HAN, J., KAMBER, M. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, 2001. ISBN 1-55860-489-8.
- [Ha05] HAN, Jiawei. *Data Mining and Database Systems* [online]. 2005 [cit. 2005-04-09]. Dostupný z: <http://www.sal.cs.uiuc.edu/~hanj/index.html>.
- [M704] ISO/IEC JTC1/SC29/WG11. *MPEG-7 Overview* [online]. Martínez, José M. version 10. Palma de Mallorca : 2004 [cit. 2005-10-22]. Dostupný z: <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>.
- [M702] Neil Day: *EE Times: MPEG-7 tackles multimedia content* [online]. California. 2001 [cit. 2005-10-27]. Dostupný z: <http://www.eet.com/story/OEG20011112S0058>.
- [Or05] Oracle Corporation. *Oracle Database Online Documentation 10g Release2* [online]. 2005 [cit. 2005-10-27]. Dostupný z: http://www.oracle.com/pls/db102/portal.portal_db.