

Денис Шестаков, denis.shestakov@utu.fi
Университет Турку

Введение в мультимедийные базы данных

Информация о конкурсе «Класс-2006»

- Данный курс занял третье место на конкурсе учебных курсов по информационному поиску «Класс 2006»
- Информацию о конкурсе «Классе-2006» можно найти на сайте компании «Яндекс», которая организовала и провела данный конкурс

- Обратите внимание на другие курсы, победившие в конкурсе:

(<http://company.yandex.ru/class/news.xml>, 28 сентября 2006г.)

1 место: «**Информационно-поисковые системы**», автор А.В. Сычев (Воронежский государственный университет)

2 место: «**Тезаурусы и онтологии**», авторский коллектив под руководством В.Д. Соловьева (Казанский государственный университет, Московский государственный университет)

3 место: «*Мультимедийные базы данных*», автор Д.А. Шестаков (Университет Турку) <--- данный курс

3 место: «**Теория информационного поиска**», автор Д.В. Ландэ (Международный Соломонов Университет, Киев)

Призовое место: «**Теория баз данных и информационного поиска**», автор Э.Э. Гасанов (Московский государственный университет)

Призовое место: «**Современные задачи теоретической информатики**», автор Ю.М. Лифшиц (Санкт-Петербургское отделение математического института им. В.А. Стеклова РАН, Санкт-Петербургский государственный университет информационных технологий, механики и оптики)

Призовое место: «**Корпусная лингвистика**», автор В. Захаров (Санкт-Петербургский государственный университет)

Призовое место: «**Поиск информации в Интернете**», автор И.М. Печищев (Пермский государственный университет)

Информация о данном курсе

Структура курса:

- Ориентировочно 12 лекций по 2 часа
- Курс разбит на 14 тем (включая «введение» (1) и последнюю тему (14) для которой нет материалов); приблизительно каждая тема – одно занятие, хотя темы 5,6,8,9 вероятно потребуют более чем двух часов на каждую
- Предполагается, что студенты подготовят 8-10 презентаций по 10-15 минут на основе прочитанных статей (список рекомендуемых статей приводится); при самостоятельном (и добросовестном 😊) прохождении курса нужно ознакомиться с несколькими статьями, взятыми из приводимого списка или самостоятельно найденными
- Вопросы и упражнения к нескольким темам (приводятся в конце соответствующих слайдов)

Информация о данном курсе

Литература:

- Данные слайды и текст лекции к теме 6
- Указывается в конце некоторых слайдов
- P. Rigaux, M. Scholl, A. Voisard. *Spatial Databases, with Application to GIS*, Morgan-Kaufmann, 2002
- Ш. Шаши, Ч. Санжей. Основы пространственных баз данных. Прентис Холл, 2004
- L. Dunckley. *Multimedia Databases – An Object-Relational Approach*, Addison-Wesley, 2003
- V. S. Subrahmanian. *Principles of Multimedia Database Systems*, Morgan Kaufmann, 1998
- S. Khoshafian, A. B. Baker. *Multimedia and Imaging Databases*, Morgan Kaufmann, 1996
- Т. Кормен, Ч. Лейзерсон, Р. Ривест, К. Штайн. *Алгоритмы. Построение и анализ*, 2-ое издание, Вильямс, 2005

Примечания и благодарности

Термины, встречающиеся в данном курсе, могут расходиться с терминами, общепринятыми в русскоязычной литературе.

Любые замечания, касающиеся некорректно переведенных терминов, а также ошибок, неточностей, опечаток и т.д., можно присылать на intro-mmdb@yandex.ru указанием номера слайда в заголовке письма.

Список найденных ошибок и корректировок будет добавлен к материалам курса.

Выражаю **признательность**:

- Компании «Яндекс» и сотрудникам «Яндекса» за проведение и организацию конкурса «Класс-2006»
- Хен Шену и Юкке Техоле за материалы, которые использовались при подготовке данного курса
- Моим коллегам за плодотворное обсуждение и ценные замечания;
- Рецензентам за конструктивную критику и полезные рекомендации

Введение в мультимедийные базы данных

1. Введение

- 1.1. Содержание курса
- 1.2. Мультимедийные типы данных
- 1.3. Пример использования мультимедийной базы данных
- 1.4. Сферы применения
- 1.5. Эволюция СУБД
- 1.6. Определение ММСУБД

Содержание курса: основные темы

К данной теме приведен текст лекции

1. Введение
2. Программное обеспечение для мультимедиа
3. Мультимедийные устройства
4. Физическое хранение мультимедийных данных
5. Организация и управление большими объектами
6. ~~Текстовые базы данных~~
7. XML и XML-базы данных
8. Структуры многомерных данных
9. Пространственные базы данных
10. Базы данных для изображений
11. Базы данных для видео
12. Аудио базы данных
13. Интеграция мультимедийных данных
14. Мультимедийные базы данных (самостоятельное знакомство с 2-3 статьями)

Наиболее сложные темы

Содержание курса

Данный курс о **базах данных** для мультимедийной информации

- Представление данных
- Структуры хранения данных
- Запросы, поиск, извлечение информации
- Индексирование

Не о:

- Использовании программных продуктов
- Средствах разработки для мультимедиа
- Создании контента
- Представлении мультимедийных данных
- Анимации

Мультимедийная революция

- Мультимедиа? Документ/набор данных, содержащий как минимум два разных медиа-формата
- Мультимедиа и компьютерная графика – необратимые тенденции
- Улучшение количества и качества информации (в сравнении с обычной текстовой информацией)
- Значительные улучшения в взаимодействии человека с компьютером
- Богатый и выразительный способ представления, просмотра и взаимодействия с информацией
- «Вторая информационная революция»
- Кардинальные изменения в бизнесе, науке, технике, производстве, медицине, искусстве, сфере развлечений и т.д.
- Принципиальная цель: вся информация в цифровом виде (до сегодняшнего дня в основном в бумажном виде)

Что делает коллекцию мультимедийных документов мультимедийной базой данных?

Мультимедийные файлы и архивы

- Простой просмотр и извлечение
- Запросы отсутствуют
- Используемое программное обеспечение: например, веб-сервер и веб-браузер
- Пример: Третьяковская галерея – экспозиция
(<http://www.tretyakovgallery.ru/russian/exposit/>)

Что делает коллекцию мультимедийных документов мультимедийной базой данных?

Мультимедийные базы данных

- Запросы по содержимому, обновления
- Управление параллельным выполнением операций, восстановление
- Используемое программное обеспечение: объектно-ориентированные или объектно-реляционные СУБД
- Пример: музей Эрмитаж – поиск QBIC по цвету и композиции

(<http://www.hermitagemuseum.org/cgi-bin/db2www/qbicSearch.mac/qbic?selLang=Russian>)

Мультимедийные типы данных

- Текст
- Графика
- Звук
- Видео

Может что-то забыли?

Пожалуй да ... 😊 , но еще не сейчас и даже не 'завтра':

- Тактильные ощущения
- Запах
- Вкус
- ...

Мультимедийные типы данных

Текст

- Присутствует в большинстве мультимедийных приложений; дополняет/поясняет нетекстовые форматы данных
- Визуальное разнообразие достигается шрифтами
- Самый компактный (при хранении) тип данных

Аудио

- Все более популярный тип данных
- Множество форматов (wav, cd, mp3, au, aiff, qt, ra, wma, ...)
- Оцифрованное аудио имеет относительно большой размер (одна секунда занимает десятки килобайт)
- Используется сжатие (коэффициент сжатия mp3 - 12:1)
- Более компактное представление аудио-данных: синтезируемая музыка в формате MIDI, MPEG-4SA(Structured Audio)

Мультимедийные типы данных

Статические растровые изображения

- Черно-белые/градации серого/цветные
- Одностраничное изображение в хорошем разрешении занимает несколько мегабайт
- Множество графических форматов (bmp, gif, tiff, jpeg, psx, png, ...)
- Формат JPEG (текущая версия - JPEG-2000): коэффициент сжатия обычно больше 10

Мультимедийные типы данных

Цифровое видео

- Последовательность кадров (фреймов) (= статичных изображений)
- Требует много дискового пространства
- Коэффициент сжатия более высокий, чем у статичных изображений (мало отличий между последовательными фреймами)
- Скорость компрессии/декомпрессии и передачи должна быть не менее 20-30 фреймов в секунду
- Анимационное видео более компактно (синтезированные изображения, использование стандартных шаблонов)
- MPEG-4: объектно-базирующееся представление, специальные методы

Мультимедийные типы данных

Векторная графика

- Двух- или трехмерные рисунки/чертежи, модели, карты
- Относительно компактный размер: преимущественно состоит из объектов, а не пикселей
- Параметры (мета) объектов: масштаб, ориентация, вращение и т. д.
- Типичные приложения: автоматизированное проектирование и изготовление чертежей, географические информационные системы (ГИСы)

Интегрированные документы (текст и изображения)

- Могут создаваться современными текстовыми редакторами

Мультимедийные типы данных

Интегрированные аудио и видео

- Структуры чередования (interleaved structures), задающие временные последовательности для воспроизведения аудио и видео-поток
- Форматы: AVI (Audio Video Interleave by Microsoft), ASF (Microsoft), QuickTime (Apple), RM (RealMedia)

Универсальные интегрированные мультимедийные презентации

- MS PowerPoint, Macromedia Director, SMIL

Гипермедиа

- Нелинейное представление мультимедийных данных (содержит гиперссылки)

Возможный сценарий (использования мультимедийной бд)

ФСКН РФ (Федеральная служба РФ по контролю за оборотом наркотиков) ведет расследование широкомасштабной торговли наркотиками на территории какого-то округа РФ

Для сбора данных используются следующие устройства:

- ✓ Камеры видео-наблюдения (записывают происходящее в определенных местах); 50-100 камер; миллионы фреймов с каждой камеры; огромное количество видео данных должно быть приведено в порядок и проиндексировано
- ✓ Устройства прослушивания телефонов (записывают телефонные разговоры в течение длительного периода времени); распознавание и организация записей – очень существенная работа

Возможный сценарий (использования мультимедийной бд)

- ✓ Фотокамеры (делающие фотографии, т.е. статичные изображения); фото-изображения должны быть оцифрованы и заархивированы для дальнейшего поиска и извлечения информации
- ✓ Помимо этого работают со следующими текстовыми данными: документы: письма, приказы, счета, соглашения и т.д.
- ✓ Структурированные реляционные данные: телефоны, адреса, имена, банковские транзакции и т.д.
- ✓ Географическая информация: карты, рельефы местности, маршруты и т.д.

**В расследовании используется вся эта информация вместе
(= мультимедиа)**

Возможные запросы

- а) *Текстовый запрос: найти все документы (из милицейских архивов, архивов газет, заявлений свидетелей, банковских транзакций), в которых подозреваемое лицо/компания косвенно или прямо совершила какие-либо операции с компанией АБВ. Документы должны индексироваться на основе семантического значения ключевых слов.*
- б) *Запрос по изображению: по имеющейся фотографии человека найти другие фотографии с этим же человеком.*
- в) *Аудио-запрос: определить говорящего по записи.*
Потребуется специальные методы обработки звука; основная идея: характеристический вектор (feature vector) – характеристика речи

Возможные запросы

- г) Видео-запрос: *найти видео сцены в которых подозреваемые совершают определенные действия (например, встречаются). Требуется специальные методы индексирования.*
- д) Простой гетерогенный (мультимедийный) запрос: *найти людей, подозреваемых в преступлении И переведших (или получивших) деньги на счета компании АБС.*
- е) Другой гетерогенный запрос: *найти людей, сфотографированных с господином Х. И подозреваемых в преступлении И переведших деньги на счета компании АБС.*
- ж) Сложный гетерогенный запрос: *найти людей, которые были в контакте с господином Х. И подозреваемых в преступлении И ...*

Сферы применения

а) Мультимедийные образовательные сервисы:

- Удаленное обучение
- Учебные материалы
- Архивы аудио-/видеоматериалов (для образования)
- Возможность предварительного просмотра

б) Видео по требованию:

- Выбор видеоматериала (фильма, ...), возможно с помощью запросов
- Возможность предварительного просмотра; перемотка вперед/назад
- Высокая пропускная способность
- Простой способ оплаты
- В ближайшем будущем, но еще не сегодня

Сферы применения

в) Экспертные системы:

- Ремонт машины: автоматические помощники в выполнении различных авто-работ; просмотр демонстрационного видео нагляднее и удобнее чтения руководств по ремонту
- Медицина: стандартные хирургические операции

г) Туристические компании (агентства путешествий):

- Интеллектуальный мультимедийный тур-агент
- Пользователь задает параметры желаемого путешествия/отдыха
- Система возвращает маршруты и места, удовлетворяющие требованиям
- Система генерирует мультимедийные презентации возможных вариантов путешествия
- Требования пользователя строго определяют содержимое презентации
- После выбора путешествия (маршрута, места), система предлагает план поездки, отели, ...
- Система запрашивает системы бронирования билетов (авиа, жел., ...), отелей и списывает получившуюся стоимость со счета пользователя

Сферы применения

д) Электронная коммерция:

- Онлайн информация о продуктах/услугах: картинки, разъяснения, наличие и т.д.
- Возможность выполнять запросы
- Удобные системы оплаты онлайн (кредитные карточки, интернет-валюты, ...)

е) Медицинские информационные системы:

- Медицинские карты (записи) пациентов, содержащие изображения с рентгеновских, ультразвуковых и т.д. обследований
- Строгая конфиденциальность информации
- Использование для диагностики, наблюдения, исследовательских целей
- Автоматизированные методы: например, методы распознавания образов

Основной фокус курса

Общие наблюдения:

- Все рассмотренные области применения имеют общие аспекты и проблемы
- Цель: найти общие «базисные» алгоритмы, не связанные с конкретной областью применения, которые можно использовать во множестве приложений лишь с незначительной модификацией
- Это позволит разработать универсальную мультимедийную СУБД (ММСУБД); возможно, ММСУБД может быть построена как расширение стандартной СУБД

В этом курсе мы сфокусируемся на:

Представлении, организации и индексировании мультимедийных данных, а также поиске по мультимедийным данным

Технологии для мультимедиа

(позволяющие использовать мультимедиа)

- Аппаратные средства («железо»): сканеры, плоттеры, цифровые камеры, видеокамеры, видео/аудио-карты, мониторы высокого разрешения, сенсорные экраны, оборудование для виртуальной реальности
Высокопроизводительные мультимедийные рабочие станции доступны и приемлемы по цене
- Широкополосные сети (WAN, LAN), оптоволокно, сетевые стандарты (Gigabit Ethernet, FDDI - интерфейс оптоволоконной передачи, ATM - асинхронный режим передачи)
- Устройства хранения большой емкости: оптические диски, многоприводные/многодисковые устройства
- Программное обеспечение для обработки изображений: сжатие (JPEG, MPEG), анализ, фильтрация, сегментация
- САПР (CAD) и анимационное программное обеспечение: двух- и трехмерная графика, применения в науке, технике, медицине, компьютерных играх и т.д.
- Распознавание образов (символов, форм, фигур): например, нейронные сети
- Усовершенствованные программные средства: объектно-ориентированные языки, объектно-ориентированные базы данных, операционные системы, многопоточность и т.д.

Эволюция СУБД

а) Сетевые и иерархические модели:

- Появились после ранних файловых систем
- Включали язык описания данных (DDL = Data Description Language) и язык манипулирования данными (DML = Data Manipulation Language)
- Поддерживали связи (отношения) один к многим
- Навигационный способ обработки
- Администратор БД (DBA) задавал (физическое) размещение, порядок и т.д. данных
- Зависимость от физического расположения данных

б) Реляционная модель:

- Логическая модель данных, независящая от физического способа хранения
- Простая и элегантная
- Математически-обоснованная теория
- SQL-запросы: говорят что извлечь, но не как извлечь

Эволюция СУБД

- в) Семантические и функциональные модели данных:
 - Лучшее приближение при моделировании окружающей реальности
 - Коммерческие системы появились в 80-х годах
- г) Расширенные и объектно-реляционные модели:
 - Включают в себя некоторые объектно-ориентированные признаки (абстрактные типы данных, идентификация объектов, хранимые процедуры, задаваемые пользователем функции)
 - SQL3
 - Поддержка мультимедиа часто ограничена большими (массивными) двоичными объектами (BLOB = Binary Large Object)
- д) Вложенные (nested) реляционные и сложно-объектные модели:
 - Непервая нормальная форма (NF^2)
 - Конструкторы: множество, кортеж (tuple), список
 - Древовидная структура
 - В основном академические прототипы

Эволюция СУБД

е) Объектно-ориентированная (ОО) модель данных:

- Идентификация объектов, абстрактные типы данных, наследуемость, графо-структурированные связи
- Многочисленные коммерческие системы с середины 80-х
- ОО модель и мультимедиа имеют много общего; тип = представление + операции

ж) Требования к модели данных для мультимедиа:

- Должна поддерживать мультимедийные типы данных (текст, изображения, звук и т.д.); не должно быть обычных файловых ссылок на оптический или иной носитель, что распространено в расширениях реляционной модели или ОО модели данных
- Должна поддерживать поиск по содержимому любого типа мультимедийных данных (например, поиск по всему тексту, или поиск заданного лица по фотографиям)

Упрощенное представление о мультимедийных бд

Популярные, но упрощенные представления о мультимедийной базе данных:

- а) CD-ROM содержащий мультимедийные данные
- б) Мультимедийная файловая система
- в) Видео/аудио по запросу: быстродействующие параллельные диски и высокоскоростная сеть
- г) Системы обработки документов и изображений:
сканирование, хранение, индексирование, и извлечение
больших объемов печатных документов

Упрощенное представление о мультимедийных бд

- д) Реляционная бд + поддержка больших двоичных объектов (BLOB): фрагментарная (кусочная) обработка массивных двоичных объектов, пользовательские функции
- е) Объектно-реляционные бд + поддержка массивных двоичных объектов: поведение (характеристики) мультимедийных объектов могут быть реализованы в СУБД; такие системы поддерживают ряд мультимедийных типов данных
- ж) Объектно-ориентированные бд + поддержка массивных двоичных объектов: аналогично е), но более явное представление сложных мультимедийных объектов; хорошо подходит для систем автоматизированного проектирования и производства (CAD/CAM)

Упрощенное представление о мультимедийных бд

- з) Графические инструментальные средства начальных этапов (graphical front-end tools): создание графических (клиентских) интерфейсов (например, форм), связанных с серверными (back-end) базами данных
- и) Пространственные бд: географические информационные системы (ГИСы); возможность пространственных запросов, пространственные отношения, поиск по пространственной близости
- к) Базы данных для систем автоматизированного проектирования и производства (CAD/CAM): двух- и трехмерные графические объекты

Ни одно из этих представлений не включает в себя все характеристики мультимедийной базы данных, хотя все перечисленные свойства уместны и существенны.

Определение ММСУБД

1. Поддерживает изображения, аудио и другие мультимедийные типы данных
2. Может оперировать с очень большим количеством мультимедийных объектов
3. Поддерживает высокопроизводительную систему хранения данных большой ёмкости: иерархическое управление запоминающими устройствами (Hierarchical Storage Management) (оперативное, полуоперативное и автономное хранение)
4. Предлагает следующие СУБД-средства: долговременное хранение, транзакции, управление параллельным выполнением операций, восстановление при отказах, запросы с декларативными конструкциями высокого уровня, контроль версий (versioning), ограничения целостности, безопасность и защита данных, высокая производительность
5. Средства информационного поиска: поиск по точному совпадению, вероятностный поиск, поиск по содержимому, ранжирование результатов

Другие важные характеристики ММСУБД

- 1) Пространственные типы данных и пространственные запросы
- 2) Интерактивные запросы, релевантная обратная связь (relevance feedback), уточнение (refinement) запросов
- 3) Автоматическое извлечение и индексирование признаков (характеристик)
- 4) Одно- и многомерное индексирование
- 5) Индексирование для ассоциативного поиска (content retrieval)
- 6) Кластеризация сложных объектов
- 7) Структуры хранения и памяти для больших двоичных объектов
- 8) Оптимизация мультимедиа-запросов

Архитектурные соображения

Текущее состояние:

- Реляционные или расширенные реляционные СУБД, с поддержкой больших двоичных объектов
- Системы иерархического управления запоминающими устройствами (Hierarchical Storage System)
- Модуль извлечения информации (поиск по содержимому документов)

Идеально:

- Расширяемая система с объектно-ориентированными возможностями
- Поддержка запросов и транзакций для мультимедийных объектов
- Поддержка сложных объектов (состоящих из мультимедийных подобъектов)

Клиент-серверная архитектура:

- Сервер выполняет стандартные СУБД-функции + мультимедийные расширения
- Клиент использует пользовательский интерфейс
- Между клиентом и сервером - интерфейс запросов высокого уровня (API)

Схематическая архитектура ММСУБД

