

Введение в мультимедийные базы данных

2. Программное обеспечение для мультимедиа

2.1. Редактирование и обработка изображений

2.2. Редактирование и обработка аудио

2.3. Редактирование и обработка видео

2.4. Обработка документов

2.5. Гипермедиа

2.6. Интегрированное программное обеспечение

2.7. SQL и мультимедиа

2.8. Метаданные

Программное обеспечение

Виды программного обеспечения:

- Средства редактирования: редактирование текста, изображений, звука, видео
- Средства разработки: например, для моделирования двух- и трехмерных графических объектов
- Программные средства конечного пользователя: клиентские приложения, включающие в себя мультимедийные элементы
- Средства для подготовки контента и презентации: например, MS PowerPoint
- Средства для работы с формами: программные интерфейсы на базе форм (form-based interfaces)

Программное обеспечение для мультимедиа можно разбить на две группы:

- Для разработки и создания мультимедийных приложений: разработка, создание, редактирование
- Для работы с мультимедийными приложениями: извлечение, потоковая передача (streaming), просмотр

Редактирование и обработка изображений

Простейшие средства:

- Рисование элементарных графических объектов: линий, прямоугольников, окружностей, и т.д.; заполнение цветом
- Пиксель-ориентированные; ~~не для работы с графическими объектами~~
- Пример: MS Paint

Редакторы векторной графики:

- Коллекция базовых графических объектов
- Объединение базовых объектов в более крупные составные объекты
- Более компактное представление в сравнении с пиксельной графикой
- Средства автоматизированного проектирования (САПР): каркасные (wireframe) модели, моделирование поверхностей
- Визуализация: вращение, масштабирование, передвижение (по объектам)

Редакторы растровой графики:

- Яркость, контраст, резкость, сглаживание, различные фильтры

Редактирование и обработка изображений

Анализ изображений:

- Характеристики изображений: текстура, яркость, цвета
- Разбиение изображения на объекты (распознавание объектов)
- Свойства объектов: размер, форма (фигура), контур

Синтез изображений:

- При помощи исходных изображений-шаблонов
- На основе численных данных (например, в компьютерной томографии)

Функции СУБД для работы с изображениями:

- Хранение больших двоичных объектов с возможностью поблочного доступа к данным
- Хранение атрибутов, извлечение шаблонов (паттернов) для индексации и поиска

Сжатие изображений:

- Различные методы: методы преобразований (JPEG, вейвлеты («всплески»)), векторное квантование (vector quantization), методы предсказания изображения (predictive methods), фрактальное сжатие и т.д.

Редактирование и обработка изображений

Технологии баз данных требуемые для графики:

1) *Иерархическое представление графических объектов (сложно-составных объектов):*

- Рекурсивное использование конструкторов для кортежей (tuples), списков, множеств и массивов
- Объектно-ориентированные бд (ООСУБД): прямое (явное) моделирование отношений «часть-целое»
- Реляционные бд (РСУБД): процесс нормализации ведет к разбиению объекта – объект представляется в виде большого количества записей в нескольких таблицах; для работы с целым объектом (например, извлечение объекта из бд) каждый раз требуются дорогостоящие операции JOIN

2) *Поддержка поведения (набора методов) графических объектов:*

- Каждый графический элемент имеет определенные свойства (или атрибуты; совокупность атрибутов объекта называется *состояние* объекта) и набор методов (*поведение*), оперирующих над свойствами объекта
- ООСУБД: напрямую поддерживают поведение объекта
- РСУБД: может быть расширена с помощью хранимых процедур и пользовательских функций

Редактирование и обработка изображений

3) Совместное использование общих подобъектов

- Требует объектной идентификации: один и тот же подобъект может быть частью несколько составных объектов (которые могут задаваться как список идентификаторов подобъектов).

4) Одновременный (параллельный) доступ к объектам

- Некоторые ООСУБД поддерживают синхронизацию check-out/check-in (изъятие объекта из хранилища с блокировкой оригинала и копирование обратно после модификации) для составных объектов

5) Контроль версий (versioning) объектов

- Особенно важно для САПР приложений

Редактирование и обработка изображений

6) *Продолжительные транзакции*

- Возможно длительное выполнение; также вложенные транзакции

7) *Работа с большими двоичными объектами (BLOBs)*

- ООСУБД и ОРСУБД (объектно-реляционные) могут иметь специальные операции (фильтры) для работы с BLOBs

8) *Множественное представление сложных объектов*

- Возможности импорта/экспорта в различные форматы
- Интероперабельность со стандартными интерфейсами (OpenDoc, OLE 2.0, CORBA, и т.д.)

Редактирование и обработка изображений

Распознавание объектов в растровых изображениях:

- Сегментация:

Нахождение потенциальных объектов для последующего распознавания, анализируются границы (контуры, границы разрывов (резких изменений) в изображении) объектов

- Векторизация:

Трансформация растрового изображения в линии, окружности, многоугольники, трехмерные сцены

- Распознавание объектов:

Лучший распознаватель – человеческий мозг;

Методы искусственного интеллекта (например, нейронные сети) разрабатываются, требуют значительных компьютерных ресурсов

Исключение: оптическое распознавание символов (OCR), в том числе рукописных – относительно зрелая технология

- Индексирование распознанных объектов и их свойств:

Необходимо для последующего поиска, позволяет выделять и объединять в группы похожие объекты

Редактирование и обработка аудио

Источник цифрового звука:

- В большинстве случаев аналоговый аудио-сигнал; исключение: синтезируемая музыка

Оцифровка (сэмплирование):

- Частота сэмплирования (дискретизации) – частота (количество сэмплов в секунду). Например, 44.1 кГц - в одну секунду записано 44100 сэмплов (дискретных значений)
- Размер сэмпла (sample size) – точность представления, количество бит на один сэмпл (например, 8, 16, 32)
- Речь: 8 кГц, 8 бит на сэмпл
- CD-аудио: 44.1 кГц, 16 бит на сэмпл

Компрессия:

- Различные методы и форматы, многие основаны на преобразовании Фурье

Редакторы аудио:

- Фильтрация, удаление шумов, частотная коррекция (equalization), реверберация (reverb), ...

Редактирование и обработка аудио

Преобразование текста в речь (процедура):

- Заменить числа и аббревиатуры полным текстом
- Определить нужные фонетические правила по словарю
- Преобразовать в фонематическую транскрипцию
- Добавить интонацию и ударения согласно правилам просодии (прим.: из словаря Даля: «просодия» - слогаударение, правильное произношение долгих и коротких слогов речи, певучесть) - сложная задача
- Преобразование в речь (воспроизведение)

Области применения речевых преобразователей:

- Аннотация документов
- Системы помощи
- Устройства для лишенных речи людей

Редактирование и обработка аудио

Распознавание речи:

- Уровни: простые команды, синтаксический анализ предложений, семантический анализ
- Методы: система обучается на большом количестве речевых образцов
- Процесс распознавания – поиск по степени близости с формами сигналов (образцов), известных системе

Хранение аудио-данных:

- Как поля в мультимедийной базе данных
- Как отдельные файлы, ссылки на которые (а также их описания) содержатся в базе данных

Редактирование и обработка аудио

Цифровой интерфейс музыкальных инструментов (MIDI) (musical instrument digital interface):

- Система кодирования для электронной (синтезируемой) музыки
- Хранится не сама звуковая информация, а наборы команд (проигрываемые ноты, ссылки на проигрываемые инструменты, значения изменяемых параметров звука) с указанием времени выполнения
- Могут использоваться 128 инструментов (у каждого свой id)
- Команды воспроизводятся по-разному, в зависимости от устройства воспроизведения
- Требуется MIDI-совместимое устройств (например, звуковая карта компьютера, синтезатор)
- Аналогия: векторная графика /MIDI/ vs. растровая графика /оцифрованное аудио/
- Гораздо более компактно чем оцифрованное аудио
- MIDI-партитуры обычно создаются с помощью секвенсеров
- MIDI-партитуры можно редактировать и микшировать с другими партиями
- Расширение и обобщение MIDI: стандарт MPEG-4 Structured Audio (MP4-SA)

Редактирование и обработка видео

Оцифровка:

- Преобразование аналогового сигнала в последовательность кадров (фреймов)
- Используются устройства видео-захвата (frame grabber) на видео-картах

Передача видео в реальном масштабе времени:

- Видеокамеры плюс оцифровка
- Высокоскоростные схемы компрессии/декомпрессии
- Сети высокой пропускной способности (требуется доставлять 30 фреймов в секунду); уменьшение размера окна и разрешения снижает нагрузку
- Потеря фреймов дает эффект прерывистости
- Области применения: видеоконференции, дистанционное обучение

Редактирование и обработка видео

Обработка видео (видеомонтаж):

- Видео-захват и воспроизведение
- Управление видеоустройствами
- У каждого фрейма есть временная метка (адрес)
- Видеофрагмент (клип) – непрерывная последовательность фреймов
- Видеоэффекты:
 - наложение графических элементов или изображений;
 - плавное изменение яркости, затухание (fading);
 - объединение;
 - монтажный переход, наплыв (transition);
 - морфинг (плавное преобразование изображения (объекта) в другое) (morphing);
 - выбор сцены, масштабирование, вращение;
 - деформация изображения (warping);
 - фильтрация (например, изменение цветов)

Редактирование и обработка видео

Области применения смонтированного видео:

- Рекламная деятельность, маркетинг
- Образование
- Развлечения

Базы данных для видео:

- Большой размер
- Большие двоичные объекты или подсистема иерархического хранения
- Индексирование:
 - свойств (признаков), задаваемых пользователем;
 - автоматически определяемых признаков
- Структурное разбиение: фрейм, видеофрагмент, сцена
- Видео по требованию: кэширование популярных видеоданных для повышения производительности

Обработка документов

Формы:

- Общепринятый способ организации информации (в бумажном и электронном виде)
- Графический интерфейс пользователя (GUI): построение запросов, редактирование
- Программные пакеты для работы с формами: легко настраиваемые клиентские средства для доступа к базам данных
- Программное обеспечение коллективной работы (groupware): электронная почта, документооборот

Редактирование форм:

- Кнопки (команды)
- Поля, перечни (заголовки, ввод данных)
- Графические (и другие мультимедийные) объекты
- Группы данных (таблицы, меню, списки)
- Привязка компонентов формы к базе данных
- Просмотр данных из бд

Обработка документов

«Построение запроса через форму» (Query by Form) - визуальное построение запроса:

- Запросы строятся непосредственно на экране путем ввода соответствующих значений или условий в поля формы
- Результаты также показываются в форме
- Аналогичен способу формулирования запроса «построение запроса по образцу» (QBE) (Query by Example)
- Может быть расширен для мультимедийных данных

Индексирование:

- Необходимо для эффективного поиска по документам
- Может быть “узким” местом для систем, поддерживающих графическое представление документов

Составные (compound) документы:

- Состоят из частей, созданных разными приложениями
- Основа: многократно используемые компоненты, возможно от разных поставщиков
- Интероперабельность: компоненты могут быть связаны с друг другом и обмениваться данными
- Части (документов) могут быть распределены по сети

Обработка документов

Интероперабельность - стандарты:

- OLE 2.0 (базируется на COM = Component Object Model/Microsoft)
- OpenDoc (базируется на SOM = System Object Model/Component Integration Labs; образован Apple, IBM, Oracle и др.)
- CORBA (Общая Архитектура Брокера Объектных Запросов - Common Object Request Broker Architecture; взаимодействие объектов в распределенной среде; поддерживается консорциумом OMG (Object Management Group))
- Web Services («заново изобретенная CORBA»); поддерживается консорциумом W3C (World Wide Web Consortium); <http://www.w3.org/2002/ws/>)

Стандартизация:

- Бинарное представление объектов
- Пользовательский интерфейс: создание, взаимодействие, представление, меню, интеграция
- Хранение: структура иерархического включения (containment hierarchy)
- Совместно-распределенные вычислительные среды

Обработка документов

Системы, поддерживающие графическое представление документов (document-imaging systems):

- Преобразование бумажной версии документа в электронную путем сканирования
- Оптическое распознавание символов (OCR) – преобразование электронного образа документа в текст; для всего документа или для определенных зон в документе
- Атрибутное индексирование (по задаваемым атрибутам) и (полнотекстовое) индексирование по всему содержимому (автоматическое)
- Поддержка автоматизированных информационных потоков (документооборот в организации)
- Аннотирование: добавление текстовых пояснений

Гипермедиа

- Гипертекст → гипермедиа (текст, аудио, видео, графика, гиперссылки между ними)
- Два вида мультимедиа:
 - Линейная (неинтерактивное представление):
например, фильм в кинотеатре
 - Нелинейная (интерактивное представление) – гипермедиа: например, Веб
- Более универсальна в сравнении с составными документами:
 - Гибкость связей между/внутри документами
 - Средства навигации/просмотра
 - Словарь терминов (глоссарий) помогает при навигации
 - Различные уровни детализации для документов
 - Системы создания гипермедиа позволяют редактировать различные типы медийных данных
- Эффективна для groupware-приложений (программное обеспечение коллективной работы)
- ‘Память организации’ (corporate memory) - совокупность данных, информации и знаний, необходимых для достижения стратегических целей организации

Гипермедиа

Виды ссылок:

- Перейти к, масштабирование, прокрутка
- Настройка гипермедиа: невидимые ссылки

Организация связей:

- Связывающие ссылки: ссылки между документами или частями документов
- Ссылки, привязанные к определенным областям (прямоугольным областям на экране)
- Ссылки, привязанные к иконкам/кускам текста
- Ссылки, вызывающие определенные действия

‘Гипермедиа’ и ‘мультимедиа’:

- Гипермедиа – навигационный механизм для мультимедиа
- Включает в себя возможности поиска по мультимедийным данным
- Различные представления и уровни детализации мультимедийных данных
- Наиболее массовый тип гипермедиа: веб-документы

Интегрированное программное обеспечение

Программные средства для подготовки контента и презентации:

- Средства презентации: редактор + средство просмотра (например, MS PowerPoint)
- Образно-графические системы создания контента (iconic authoring systems): интерактивное мультимедиа
- Гипермедиа: нелинейное представление, основа – гиперссылки
- Системы с временными шкалами (timeline-based systems): расположение мультимедийных компонентов вдоль временной шкалы

Хранение мультимедиа:

- Большие двоичные объекты или структурные объекты
- Объектно-ориентированные базы данных: сложно-составные объекты с совместно используемыми подобъектами
- Структура и метаданные мультимедийных презентаций
- Поддержка строго ограниченных по времени чередующихся презентаций

Операционные системы:

- Общего назначения (Windows, Linux) или специализированные (BeOS)

Интегрированное программное обеспечение

Графический интерфейс пользователя (GUI):

- Разработан преимущественно в 1970-ые годы
- Впервые использован в Apple Macintosh
- Визуализация: действия на компьютере выполняются при помощи манипуляций с графическими объектами (иконками)
- Реализация 'рабочего стола': папки, календарь, корзина, ...
- Объектно-ориентированная направленность:
 - Выполнение операций связано с графическими объектами на экране
 - Объекты принадлежат к разным классам (находятся в одной папке, документы Word и т.д.)
- Шаблоны для последовательности выполняемых операций

Интегрированное программное обеспечение

Функции мультимедийных баз данных необходимые для приложений:

- Сопровождение (поддержка) мультимедийных объектов
- Поддержка метаданных
- Совместное использование данных несколькими пользователями
- Одновременные (параллельные) обновления; управление транзакциями
- Просмотр, редактирование, выполнение запросов и поиск по мультимедийным объектам
- Навигационный доступ и интерактивное уточнение запросов
- Атрибутное индексирование и (полнотекстовое) индексирование по всему содержимому

SQL и мультимедиа

- Стандарты SQL: 1986, 89, 92 (SQL2), 99 (SQL3), 2003 (SQL2003) (<http://www.wiscorp.com/SQLStandards.html>)
- В SQL3 добавлены некоторые объектно-ориентированные возможности, хранимые процедуры, пользовательские (определяемые пользователем) функции, спецификация формулировки запросов (API для (реляционных) систем баз данных, поддерживающих SQL
- В SQL2003 добавлена поддержка XML, функции с переменным окном (window functions), генерация последовательных значений и т.д.
- Типы данных для больших (массивных) объектов:
 - BLOB (Binary Large Object)
 - CLOB (Character ...)
 - NCLOB (National Character ..)
 - BFILE (указатель на бинарный файл вне бд; только для чтения)
- Взаимодействие с большими объектами не стандартизировано
- Расширяемость за счет хранимых процедур и пользовательских функции
- Производитель-зависимые расширения: например, пакет DBMS_LOB от Oracle: операции для доступа и работы с фрагментами больших объектов или с целыми объектами

SQL и мультимедиа

Другие способы построения запросов к мультимедийным данным:

- Поиск по атрибутам (традиционный способ):
Фиксированный набор атрибутов, индексирование (например, с помощью B+-деревьев)
- Поиск по тексту:
Мультимедийные объекты дополняются ключевыми словами (тэгами), задаваемыми вручную, и текстовыми аннотациями. Поиск - полнотекстовое сканирование, инвертированные файлы, основанные на ключевых словах, сигнатурные файлы и т.д.
- Поиск по всему содержимому:
Автоматическое определение/извлечение характеристик (свойств) мультимедийных объектов. Поиск – указанные в запросе свойства сравниваются (на степень схожести) со свойствами мультимедийных объектов в коллекции.

Метаданные

Метаданные – данные о данных:

- Важная часть любой бд
- Описание каждого объекта в коллекции
- Необходимы для организации и управления хранимыми объектами

Онтология (предметной области) определяет понятия (концепты), их значения и их взаимосвязи в предметной области. Состоит из:

- Списка терминов и их определений
- Списка типов (видов) объектов
- Связей между понятиями и объектами

Метаданные

Классификация метаданных:

- Независящие от содержания (например, для фотографии – имя фотографа, время съемки)
- Описывающие содержание (текстовая аннотация):
 - предметно-зависимые метаданные (относящиеся к предметной области);
 - предметно-независимые метаданные
- Зависящие от содержания (характеристики, извлеченные из мультимедийных объектов)

Примеры:

Медиа	Незав.от содерж.	Описывающ.содерж.	Завис.от содерж.
Текст	Дата обновления, расположение	Ключевые слова, язык, формат	Границы подразделов в тексте
Речь	Время начала и конца, степень достоверности распознавания	Имя говорящего (-их)	Просодические метки, образцы (характеризующие говорящего)
Изображение	Автор, название, время	Ключевые слова, формат	Хар-ки изображения (гистограмма, сегментация, ...)
Видео	Название, время, распространитель	Параметры съемки (расстояние, крупный план, ...)	Границы кадров, хар-ки фреймов (гистограмма, освещенность, ...)

Метаданные

Стандартизация:

- ISO/IEC 11179: стандарт для представления метаданных для организаций
- Dublin Core Metadata Element Set: для описания различных информационных ресурсов
- Resource Description Framework (RDF): метамодель данных на основе XML; универсальный формат для веб-данных; поддерживается консорциумом W3C
- MPEG-7 (Multimedia Content Description Standard, ISO/IEC стандарт): для описания содержания мультимедийных данных, основан на XML