

多媒体技术

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题

1.1 概述

产生原因:信息量的增加,媒体种类的增加,对信息的管理和检索变得越来越困难。

在多媒体数据库系统中,除了文本和其他的离散数据外,视频和音频信息也将被存储、处理和检索。为了支持这些功能,多媒体数据库系统需要适当地存储机制和文件系统。

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题

1.1 概述

数据库系统的一个重要概念是数据的独立性——用户对数据的任何操作(查询、修改等)不再是通过应用程序直接进行,而必须通过向数据库管理系统发请求实现。

依据独立性原则,DBMS一般按层次被划分成三种模式:物理模式、概念模式、外部模式。

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题

1.1 概述

物理模式的主要职能是定义数据的存储组织方式。如数据库的文件格式、索引文件的组织方法,数据库在网络上的分布方式等。

概念模式:定义抽象现实世界的方式。概念模式通过数据模型来描述,数据库系统的性能与数据模型直接相关。数据库数据模型先后经历了网状模型、层次模型、关系模型等阶段。关系模型因为具有比较完整的理论基础,在目前的数据库中居主导地位。

外部模式:指概念模式对用户有用的那一部分。

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题

1.1 概述

关系模型:把现实世界事物的特性抽象成数字或字符串等表示的属性,每一种属性都有固定的取值范围。这样,每一个事物都有一个属性集及对应其属性的值集合。

多媒体数据在数据库管理中存在的问题:数据不规则,没有一致的取值范围,没有相同的数据量级,缺少相似的属性集。

传统数据库可以在用户给出查询条件后迅速地检索到正确的信息,多媒体数据如何检索?如何表达多媒体信息的内容?

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题

1.2 多媒体数据管理问题

1.2.1 传统的数据管理

DBMS: Database Management System: 让设计者结构化的表示信息,让用户查询或修改这些信息,让这些数据便于管理,并且还能够管理对数据的并行操作。

DBMS的主要组成部分:

📖 查询处理器

📖 事务管理器

📖 存储管理器

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.1 传统的数据管理

DBMS除提供存储管理功能外，还有其他特性：

数据的持久性：数据的生存期应超过处理程序和技术生存期。例如：在很多应用中数据必须保存在数据库中几十年甚至更长时间。而在这期间，计算机技术不断发展，操作系统在改变，其他程序也在改变，这就要求DBMS在外围的程序发生改变时，仍然能够对数据进行处理。

数据视图的一致性：在多用户系统中，在特定的时候处理数据库查询时，系统提供数据视图的一致性是非常重要的。

7

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.1 传统的数据管理

DBMS除提供存储管理功能外，还有其他特性：

数据的安全性：在系统出现错误的时候，数据库的安全性和数据库的整体防护性是对DBMS最重要的要求之一。这个属性通过使用事务概念来提供。

数据查询和检索：在数据库中存储了许多不同的信息，他们能通过数据库查询被检索到。数据库查询已被公式化为查询语言。另外，在数据库中的每个条目及它的状态信息（如条目已被修改过）需要通过适当地检索来提供。

8

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

(1) 数据量巨大且媒体之间量的差异也极大，从而影响数据库的组织和存储方法。只有组织好多媒体数据库中的数据，选择设计好合适的物理结构和逻辑结构，才能保证磁盘的充分利用和应用的快速存取。另外，数据可能需要分布在网络环境之中，这也会为存取带来问题。

关系型数据库中，BLOB（Binary Large Object）类型的结构化问题：BLOB是数据库系统的多媒体信息存储类型。但BLOB本身不支持结构化。

9

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

(2) 媒体种类的增多增加了数据处理的困难。每一种多媒体数据类型都有自己的一组最基本的操作，适当的数据结构和存取方法。虽然我们从媒体数据的特性来分可以把它分成几个大类，但是在具体的应用中会有很多类型出现，对数据库管理系统来说，应该支持这些类型。

10

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

(3) 数据库的多解查询：传统的数据库查询只处理精确的概念和查询。在多媒体数据库中非精确匹配和相似性查询将占相当大的比重。因为即使是同一个对象用不同的媒体进行表示，对计算机来说差别也很大。另外，对纹理、颜色、形状等本身就不易于精确描述的概念，如果在对图像、视频进行查询时使用，也只能是一种模糊的非精确的匹配方式。

媒体的复合、分散、时序性质及其形象化的特点，注定要使数据库不再是只通过字符进行查询，而是应该通过媒体的语义进行查询。然而，我们却很难了解并且正确处理许多媒体的语义信息。

11

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

(4) 用户接口的支持：多媒体数据库的用户接口无法用一个表格来描述，对于媒体的公共性质和每一种媒体的特殊性质，都要在用户的接口上、在查询的过程中加以体现。例如，在查询的过程中，很多情况下，用户不知道该如何来描述自己的查询要求，所以，多媒体数据库对用户的接口要求不仅仅是接收用户的描述，还要协助用户。另外，多媒体数据库的查询结果将不仅仅是传统的表格，而将是丰富的多媒体信息的表现。

例如，描述性的搜索方法。示例型的搜索方法。

12

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

- (5) 多媒体信息的分布对媒体数据库带来了巨大的影响。

这里所说的分布，主要是指以WWW为基础的分布。Internet网的迅速发展，网络上资源的丰富，多媒体数据库系统将来肯定要考虑如何从WWW网络信息空间中寻找信息，查询所要的数据。

13

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

- (6) 长事务：数据库中事务的概念：事务的四个特性ACID。
A(Atomicity)：原子性。不可再分。
C(Consistency)：一致性。
I(Isolation)：隔离性。中间结果不允许他人看到或操作。
D(Duration)：持续性。事务一旦发生，其结果持久保存。

基于这些特性，事务都比较短小。但是，在多媒体数据库中的某些场合，短事务不能满足需要（例如视频数据的读取，可能需要持续几个小时），因此应该增加处理长事务的能力。

14

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

- (7) 服务质量的要求：许多应用对多媒体数据的传输、表现和存储的质量要求是不一样的。系统所能提供的资源也要根据系统运行的情况进行控制。

对每一类多媒体数据都必须考虑这些问题：如何按所需要的形式实时地、逼真地表现数据？当系统不能满足全部的服务要求时，如何合理地降低服务质量？能否插入和预测一些数据？能否拒绝新的服务请求或撤消旧的服务请求？

15

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

- (8) 多媒体数据管理还要考虑版本的控制问题：在具体的应用中，往往涉及对某个处理的不同版本的记录和处理。版本的含义：历史版本或选择版本。

需解决多版本的标识和存储、更新和查询，尽可能减少各版本所占存储空间，控制版本访问权限。

16

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.2 多媒体数据管理问题
- 1.2.2 多媒体带来的问题

多媒体对数据库设计的影响主要表现在以下几个方面：

- (9) 数据实时传输：连续数据的读和写操作必须实时完成，连续数据的传输应优先于其他数据库的管理行为。

(10) 信息重组：应支持复合媒体在各通道分离后存入数据库。例如，将Video分解为影像、配音等信息，把这些信息分别存储到数据库中，必要时各种分离的信息能重新组织后输出。

17

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.3 多媒体数据与数据库管理

在数据库中常用的多媒体数据有：

静态：字符，数值，文本，图形，图像等静态的数据

动态：声音、视频、动画等

18

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.3 多媒体数据与数据库管理

字符数值：简单，规范，易于管理，传统数据库主要是针对这种数据的。在多媒体数据库中仍然需要管理大量这种类型的数据。

文本数据：从计算机内部来看，文本数据是由一个具有特定意义的字符串表示的。自然语言的理解技术的不成熟使得查询文本数据难度加大。在通用形数据库系统中，对文本数据处理较为常见的检索操作一般采用关键词检索和全文检索等方法。

关键词检索：自动或手工生成能反映该文本数据主题的关键词的集合。并将其存储在数据库中，检索时通过某些关键词的匹配找到所需的文本数据。全文检索支持根据文本数据中的任何单词或词组检索。

19

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.3 多媒体数据与数据库管理

声音数据：分为两类。

用符号表示的，数据量小，对它的存储和查询可采用类似于文本数据的方法。

以波形表示的数据，占用的空间比较大，检索也十分困难。

20

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.3 多媒体数据与数据库管理

图形数据：图形数据的数据库管理已有一些成功的应用范例，例如地理信息系统，工业图纸数据库等。由于图形数据可以分解为图元，所以只要有合理的描述模型来描述这个分层的结构，但管理方法和检索使用都需要明确的应用背景。

图像数据：目前在图像领域已有很多研究：属性描述，特征提取，分割，纹理识别，颜色检索等。对专业性很强的应用，已有一些比较成功的研究成果。对多媒体数据库来说，将更强调对通用图像数据的管理和查询。

21

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.3 多媒体数据与数据库管理

视频数据：时间属性带来了新的复杂程度。

例如，在检索和查询操作中，检索和查询的内容可以包括镜头、场景、内容等许多方面，这些传统的数据库中是不存在地问题。对于基于时间的媒体来说，为了真实的再现就必须做到实时，并且必须考虑和其他媒体的合成和同步。

22

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.4 对数据的操作

在数据库系统中，对每个媒体可能有不同类型的操作：输入、输出、修改、删除，比较和求值。

输入操作：数据将被写入数据库中，在原始的数据之后，可能需要附上描述性数据。如果在对音频和视频信息输入操作过程中，还需要MDBMS为之选择合适的服务器和磁盘。

输出操作：是从数据库中根据记录数据读原始数据。

修改操作：通常是对原始数据进行的。

23

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题
- 1.4 对数据的操作

删除操作：在数据删除操作期间，必须保持数据的一致性。例如，一个条目的原始数据被删除了，所有依赖于这个原始数据的其他数据类型也将被删除。

求值操作：对原始数据和记录数据进行求值的目标是产生相关的描述性数据。例如，当要求对传真文档进行存储时，可以使用字符识别软件（OCR）进行处理。

24

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题

1.4 对数据的操作

查询操作：许多对MDBMS的查询包含查找和检索已存储的数据。这些查询基于信息的比较。

查询操作的一般实现方法：把特定媒体中的单个模板和存储的原始数据进行对比，这种搜索不是非常有效。

另一种方式：模式识别，这时对原始数据进行处理。这种方法一般用在特定的系统中。

25

第八章 多媒体数据库与基于内容的检索

1. 多媒体数据库的管理问题

1.4 对数据的操作

另外，我们也可以基于面向内容的描述性数据进行比较。

例如，用户输入由几个单词组成的名词性短语，MDBMS把输入转化为谓词。这样，同义词能被系统使用和管理。这种方法允许进行内容相关的搜索，它可以用来搜索图像，也可以搜索其他各类型的媒体。

26

第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.1 多媒体数据库的一般结构形式

目前，缺少标准的多媒体数据库体系结构。已有的多媒体数据库体系结构：

□联邦型结构

□集中统一型结构

□客户/服务器型结构

□超媒体结构

27

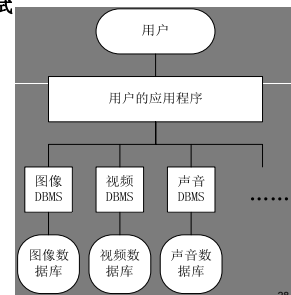
第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.1 多媒体数据库的一般结构形式

2.1.1 联邦型结构

针对各种媒体单独建立数据库，每一种媒体的数据库都有自己独立的数据库管理系统。虽然他们是相互独立的，但是可以通过相互通信进行协调和执行相应的操作。用户即可以对单一的多媒体数据库进行访问，也可以对多个媒体数据库进行访问以达到对多媒体数据进行存取的目的。



28

第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.1 多媒体数据库的一般结构形式

2.1.1 联邦型结构

特点：对多媒体数据库的管理是分开进行的，可以利用现在地研究成果直接进行“组装”，每一种媒体数据库的设计也不用考虑与其他媒体的区别和协调。

缺点：对不同类型媒体的联合操作实际上是交给用户去完成了。使得对多种媒体的联合操作、合成处理、概念查询等较难完成。

29

第八章 多媒体数据库与基于内容的检索

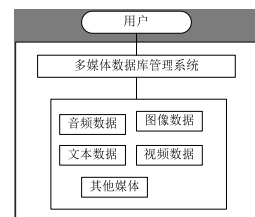
2. 多媒体数据库体系结构

2.1 多媒体数据库的一般结构形式

2.1.2 集中统一型结构

只存在一个单一的多媒体数据库和单一的多媒体数据库管理系统。

各种媒体被统一地建模，对这种媒体的管理与操纵被集中到一个数据库管理系统中，各种用户的需求被统一到一个多媒体用户接口上。



30

第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.1 多媒体数据库的一般结构形式

2.1.2 集中统一型结构

优点：建模统一、管理/操作统一、用户接口统一、查询和检索结果统一表示。

理论上，能够充分地做到对多媒体数据进行有效的管理和使用。但实际上这种多媒体数据库系统是很难实现的。目前还没有一个比较恰当而且效率很高的方法来管理所有的多媒体数据。虽然面向对象的方法为建立这样的系统带来了一定的可行性，但要真正做到还有一定的距离。

31

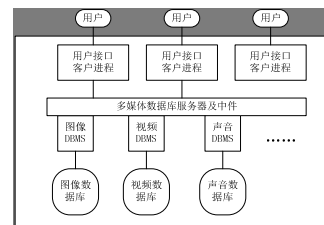
第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.1 多媒体数据库的一般结构形式

2.1.3 客户/服务型结构

减少集中统一型多媒体数据库系统复杂性的一个很有效的办法是采用客户/服务器结构。



32

第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.1 多媒体数据库的一般结构形式

2.1.3 客户/服务型结构

不同数据类型的媒体数据仍然相互独立，系统将每一种媒体的管理与操纵各用一个服务器来实现。

所有服务器的综合和操纵也用一个服务器来完成。与用户的接口采用客户进程来实现。

优点：可以针对不同的需求采用不同的服务器、客户进程组合，所以很容易符合应用的需求，对每一种媒体也可以采用与这种媒体相符合的处理办法。同时，这种体系结构也很容易扩展到为了环境下工作。但采用这种体系结构必须对服务器和客户进行仔细的规划和统一考虑，采用标准化和开放的接口界面。

33

第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.1 多媒体数据库的一般结构形式

2.1.4 超媒体型结构

在超媒体型结构中，强调对数据时空索引的组织。在这种结构的理论下，世界上所有计算机中的信息和其他系统中的信息都应该连接在一起，而且信息也要能够随意扩展和访问。

因此，没有必要建立一个统一的多媒体数据库系统，而是把数据库分散到网络上，设计合理的访问工具来访问这些信息。

另外，在超媒体的数据模型上，要通过超链建立起各种数据的时空关系，以增加检索和访问的能力。

目前，语义Web就可以看成是这种超媒体型结构的一个雏形。

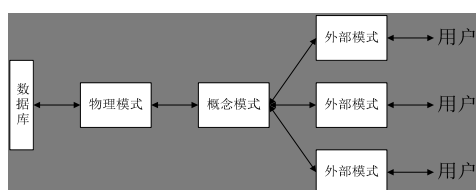
34

第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.2 多媒体数据库的层次结构

2.2.1 传统数据库的层次



35

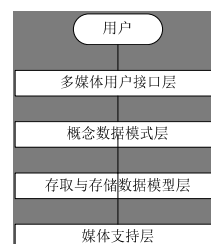
第八章 多媒体数据库与基于内容的检索

2. 多媒体数据库体系结构

2.2 多媒体数据库的层次结构

2.2.2 多媒体数据库的层次划分

总的思想：从最低层增加对多媒体数据的控制与支持，在最高层支持多媒体的综合表现和用户的查询描述。在中间增加对多媒体数据的关联和超链的处理。



36

第八章 多媒体数据库与基于内容的检索

- 2. 多媒体数据库体系结构
- 2.2 多媒体数据库的层次结构
- 2.2.2 多媒体数据库的层次划分

媒体支持层：建立在多媒体操作系统之上，针对各种媒体的特殊性质，在该层中要对媒体进行相应的分割、识别、变换等操作，并确定物理存储的位置和方法，以实现对各种媒体的最基本的数据管理和操纵。

由于媒体的性质差别很大，对于媒体的支持一般都分别对待，在操作系统的辅助下对不同的媒体实施不同的处理，完成数据库的基本操作。

第八章 多媒体数据库与基于内容的检索

- 2. 多媒体数据库体系结构
- 2.2 多媒体数据库的层次结构
- 2.2.2 多媒体数据库的层次划分

存取与存储数据模型层：完成多媒体数据的逻辑存储与存取。在该层中，各种多媒体数据的逻辑位置安排、相互的内容关联、特征与数据的关系以及超链的建立等都需要通过合适的存取与存储数据模型进行描述。

概念数据模型层：对现实世界用多媒体数据信息进行描述；也是多媒体数据库在全局概念下的一个整体视图。在该层中，通过概念数据模型为上层的用户接口、下层的多媒体数据存储和存取建立起一个在逻辑上统一的通道。

第八章 多媒体数据库与基于内容的检索

- 2. 多媒体数据库体系结构
- 2.2 多媒体数据库的层次结构
- 2.2.2 多媒体数据库的层次划分

多媒体用户接口层：完成用户对多媒体数据的查询描述和得到多媒体信息的查询结果。

这一层在传统数据库中是非常简单的，但在多媒体数据库中这一层是一个非常重要的环节。首先，必须提供某种方法使用户能够把他的要求描述出来；其次，查询和检索到的结果需要按用户的需求进行多媒体化的表现，这是在传统的关系型数据库中采用表格的形式所不能做到的。

第八章 多媒体数据库与基于内容的检索

- 2. 多媒体数据库体系结构
- 2.2 多媒体数据库的层次结构
- 2.2.2 多媒体数据库的层次划分

小结：

这里给出的多媒体数据库的层次划分是非常概念化的。多媒体数据库的结构应该能够包含像图像数据库、视频数据库、全文数据库等专业数据库类型，并且能够统一地管理和使用。

第八章 多媒体数据库与基于内容的检索

- 3. 多媒体数据模型

数据模型由3种基本要素组成：数据对象类型的集合、操作的集合、通用完整性规则的集合。

数据对象类型的集合：描述了数据库的构造。

操作的集合：数据库的运算体系。

通用完整性规则：给出一般性的语义约束。

多媒体数据库的数据模型较复杂，不同的媒体有不同的要求，不同的结构有不同的建模方法。

第八章 多媒体数据库与基于内容的检索

- 3. 多媒体数据模型
- 3.1 NF²数据模型

NF²(Non First Normal Form)数据模型是在关系型数据模型的基础上，通过扩展来提高关系型数据库处理多媒体数据的能力。

在传统的关系数据库中，表中不能有表。但由于多媒体数据库中具有多种多样的数据，这些媒体数据又要统一地在关系表中加以表现和处理，就不能不打破关系型数据库中的要求，允许表中可以有表，这就是NF²方法。

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.1 NF²数据模型

主要的手段：在关系数据库中引入抽象的数据类型，使得用户能够定义和表示多媒体信息对象。数据类型定义所必需的数据表示和操作，可以用关系型数据库语言也可以用通用的程序语言来记述。简单的说，这种模型建立在关系型数据库模型之上，就是为了继承关系数据库的成果和方法，易于实现。

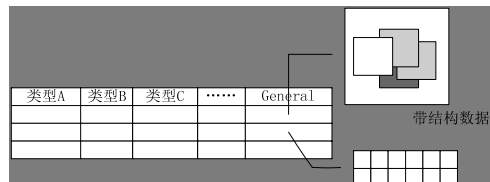
目前较多多媒体数据库的解决方法都是在关系型数据库基础上做扩展，并且在处理这些特殊的字段时自动地与相应的处理过程联系，解决了一部分多媒体数据扩展的需求。

43

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.1 NF²数据模型



44

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.1 NF²数据模型

这种方法虽然可以利用关系型数据库特有的优势，但缺点也是十分明显的，具有很大的局限性。

局限性：描述能力不够，数据模型不是很适合于多媒体应用。反映多媒体成分之间的空间关系、时间关系和媒体对象的处理方法方面仍有很多困难。在特殊媒体的基于内容查询方面、存储效率方面等都有很大的困难。

45

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.2 面向对象数据模型

面向对象数据库在数据库领域日益被重视。面向对象数据库并不等于多媒体数据库，它们在很多方面研究的侧重点不同。

可以利用面向对象数据库的一些特性来进行多媒体数据模型的建模。

46

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.2 面向对象数据模型

3.2.1 对象、属性、方法、消息

对象、属性、方法、消息：在面向对象的系统中，现实世界中所有的概念实体被模型化为对象。

组成对象的数据称为属性。

定义在对象属性上的一组操作称为对象的方法。

对象间通信和请求对象完成某种处理工作是通过消息传送实现。消息传送相当于一个间接的过程调用。对象对它能接受的每一个消息有一个相应的方法解释消息的内容，执行消息指示的处理操作。

47

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.2 面向对象数据模型

3.2.2 语义关联的描述

在多媒体数据模型中，常用的语义关联主要有以下一些：

(1) 聚集关联 (Aggregation association)：定义一个实体类的一组属性；

(2) 概括关联 (Generalization association)：继承性关系；

(3) 相互作用关联 (Interaction association)：类似于E-R模型中的实体间的Relation关系，用来表示两个实体类之间的相互作用或关系。相互作用关联定义的关系可以是一对一、一对多或多对多的关系；关联可以由用户命名，也可以带有自己的属性、操作与约束规则；

48

第八章 多媒体数据库与基于内容的检索

- 3. 多媒体数据模型
- 3.2 面向对象数据模型
- 3.2.2 语义关联的描述

在多媒体数据模型中，常用的语义关联主要有以下一些：

- (4) 示例关联：IS_INST_OF表示一个具体对象与所属实体类之间的关系，用来对具体对象建模；
- (5) has_method和has_rule关联：可以看作较为特殊的聚集关系，表示一个实体类具有数据类型为METHOD或RULE的属性；

第八章 多媒体数据库与基于内容的检索

- 3. 多媒体数据模型
- 3.2 面向对象数据模型
- 3.2.3 运算体系

在数据库系统中运算基本上有三种：定义，查询，操纵。

对采用面向对象模型的数据库而言，应该对类和对象分别定义这三种运算。

第八章 多媒体数据库与基于内容的检索

- 3. 多媒体数据模型
- 3.2 面向对象数据模型
- 3.2.3 运算体系

定义：

定义包括类的创建和对象的创建两部分。

类的创建：类标识、一组相关属性、一组操作程序、一组语义完整的约束条件和可以继承的超类集合。

对象的创建：对象创建时，对象内容与对象所属类的属性必须匹配并符合类定义的约束条件。

第八章 多媒体数据库与基于内容的检索

- 3. 多媒体数据模型
- 3.2 面向对象数据模型
- 3.2.3 运算体系

查询：查询是使用数据库的基本方法。

包括通过类名查询类结构、通过对象名或对象标识查询对象或对象的属性值、通过类名查询该类中满足某些约束条件的对象或对象的属性、对对象操作的查询。在多媒体数据库中，查询还应该包括基于内容或概念的检索等。

第八章 多媒体数据库与基于内容的检索

- 3. 多媒体数据模型
- 3.2 面向对象数据模型
- 3.2.3 运算体系

操纵：操纵运算包括插入、删除和修改。其中每种都有类和对象两种。

类的修改包括对类描述中属性集合、操作集合、约束条件集合、超类集合中元素的更新以及整个类的删除等。这些修改可能会涉及到相应的对象的修改。

对象的修改必须满足其所属类的约束条件，对象的删除也将引起其所有属性的删除。

第八章 多媒体数据库与基于内容的检索

- 3. 多媒体数据模型
- 3.2 面向对象数据模型
- 3.2.4 面向对象多媒体数据模型的特点

聚集层次：聚集抽象机制实现了不同媒体信息的有机组合，利用这种机制可以方便地得到复杂对象的层次结构以及同一级上各子对象间的顺序关系。这一点多媒体数据的存储是很有意义的。

关系建模：反映多媒体数据之间复杂的交叉关系；

方法管理：对多媒体数据的存取需要通过具体媒体技术来加以解决，例如对图像的分割、识别、压缩等。

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.2 面向对象数据模型

3.2.4 面向对象多媒体数据模型的特点

多种属性的支持和语义信息的维护：这一点实际上是要求多媒体数据库系统支持抽象数据类型。保持多媒体数据语义信息的一种良好的方式就是构造包含语义信息的类属性及操作方法的类结构。

支持特性传播：多媒体应用不仅需要聚集层次上子类能够自动继承超类的内容，同时也需要聚集层次上各类节点的某些属性能够沿着层次结构自动向下传播。

55

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.2 面向对象数据模型

3.2.4 面向对象多媒体数据模型的特点

支持模式的演进：能够动态地修改层次各节点的模式，支持模式演进。

版本管理：能够对每一次的操作都有所反映，版本的创建和控制以及版本变化的通知是十分重要的。

并发控制和数据恢复：支持多个用户对同一数据的访问。

数据共享：具有更加灵活的读写控制。

快速存储和灵活交付：应该对多媒体数据的物理表示和逻辑表示进行有效的映射管理，拥有高效的存储管理子系统，保证媒体数据流的灵活交付。

56

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.2 面向对象数据模型

3.2.4 面向对象多媒体数据模型的特点

用面向对象的方法对多媒体数据库进行建模的优点：

封装许多媒体类型通过一个公共的界面进行访问和操纵，因此即使系统发生演变，媒体的操纵仍能保持一致。

继承能够有效地减少媒体数据的冗余存储，同时它也是聚集分层和特性传播的基本方法。对象类与实例的概念有效地维护了多媒体数据的语义信息，也为聚集抽象提供了一种可行的方案。

复合对象根据复合引用的定义，对象间的引用只是被引用对象的标志符放在引用对象的属性中，从而实现共享引用、依赖引用和独立引用。为多媒体数据的关系表示提供了一种良好的机制。

57

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.3 超媒体数据模型

超媒体模型的基本结构是网状的，是由节点和链组成的有向图。

超媒体数据模型一般来说比数据库数据模型要高一个层次，它承担着建立超媒体超链联系的任务。在多媒体数据库中使用超媒体数据模型是为了建立多媒体数据之间的联系，包括时间、空间、位置、内容的关联，支持信息节点网的开放性，支持对信息结构的建模，支持浏览和搜索等新的操作。

58

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

多媒体数据库另一个极为重要的方面是用户接口。传统的数据库绝大部分工作是花在数据模型和体系结构的建立方面，对用户接口工作并不重视。但是，数据库系统为了达到最终的实用阶段，用户接口方面也要有足够的重视。

多媒体数据库的用户接口包括两个方面内容：

输入：如何将用户的请求转变为系统所能识别的形式并输入进系统成为系统的动作；

输出：如何将系统查询的得到的结果按照要求进行表现。

59

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

字符数字型接口：关键字描述、自然语言描述

关键字描述查询：

例1：带关键词的图像数据库接口

SELECT image

FROM imagesDB

WHERE keyword=“日出”

60

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

字符数字型接口：关键字描述、自然语言描述

关键字描述查询：

例2: 对图形的查询

SELECT Room

FROM Hospital_Buiding

WHERE Area(Room) > 100

61

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

字符数字型接口：关键字描述、自然语言描述

数据库自然语言接口 (NLIDB: Natural Language Interface of Database)

国外20世纪60年代开始研究, 80年代最为活跃, 20世纪90年代研究热度减弱。

中文NLIDB的研究始于20世纪70年代末期。由于中文的特殊性, 一些理论不能直接移植。

62

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

字符数字型接口：关键字描述、自然语言描述

自然语言查询存在的问题：自然语言如何进行分解以取得准确的语义(Mapping Natural Language into SQL)。

63

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

实例：

供应商S (sno, city), 工程J (jno, city),
零件P (pno,name), 供应关系SPJ(sno,jno,pno,num)
查询：帮我找一下广州的供应商提供给上海的工程的零件

找词链：找-广州-供应商-提供-上海-工程-零件

生成语义结构：属性值有“广州”，“上海”

语义可能性计算：“广州”修饰“供应商”，“上海”修饰“工程”

```
select P*  
From S,PJ,SPJ  
Where .....
```

64

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

字符数字型接口：关键字描述、自然语言描述

自然语言查询：基于自然语言的方法建立多媒体数据库的用户接口具有的特点：

(1) 共同性：对人或对象系统的变化容许度高, 可以作为不同系统的统一接口。

(2) 抽象性：用自然语言描述问题, 可以不考虑专业领域的专业知识。

(3) 模糊性：当用户的要求以及问题本身就不明确时, 用自然语言记述能反映这种模糊性。

65

第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

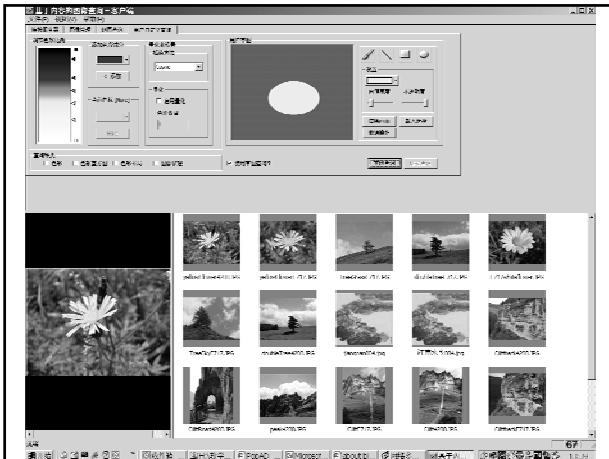
示例型接口：很多媒体的特征很难用字符表示, 这时就需要示例查询。示例的种类：

(1) 文本示例：和传统的查询类似

(2) 图像示例：给出类似图像、手绘图像的轮廓, 给出大致对象形状等。可以决定整副图像的全局特征, 也可以是该图像中某一对象的特征。

(3) 声音示例：对声音媒体的示例与图像不太一样, 因为声音无法用形象的方法显示出来。可以通过指定已有的声音示例源, 或是给出临时的声音样本。

66



第八章 多媒体数据库与基于内容的检索

3. 多媒体数据模型

3.4 多媒体数据库的用户接口

示例型接口：很多媒体的特征很难用字符表示，这时就需要示例查询。示例的种类：

(4) 视频示例：视频的示例可以分解为对图像的示例。

(5) 结构示例：结构示例指的是对信息结构的说明。例如：一个超媒体子网，用虚框指定一个结构部分，可以表示一种语义：“查找具有这种语义结构的子网”。这种结构要用语言说明会非常烦琐。

(6) 混合示例：把多种示例方法混合起来使用。

68

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.1 多媒体与媒体语义

各种多媒体数据中含有大量信息，数据和信息不同。信息来源于数据，但需要从数据中抽取。

数据可以有多种形式或类型，这些形式取决于其来源，例如图像、视频、声音或其它媒体类型。信息与任务有关，它需要根据特定的上下文并利用特定领域的知识才能推导解释出来。

同样的数据可以提供不同信息，这取决于使用者的知识和解释，以及对上下文的了解。

69

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.1 多媒体与媒体语义

信息的抽象在理解上扮演了一个非常重要的角色。早期的应用中，这个抽象的过程都是由人自己将所获取的数据转换成某种符号形式，但在其中也已损耗了相当多的信息。

抽象过程的复杂程度与任务有关，会有多种不同层次的抽象。一般来说，越接近信息原始状态的数据抽象程度越低，而符号化后的数据抽象程度就高，数据量也随着抽象程度的增加而逐步减少。抽象的过程实际上就是引入语义的过程，也与任务和领域密切相关。

目前，大多数多媒体应用是把多媒体看成一种界面工具，而并没有在各种媒体的内容上建立起联系，并且依据这些联系组织、处理和使用这些信息。

70

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.1 多媒体与媒体语义

必须要有相应的方法和工具，对多媒体的数据按不同的形式和来源获取、增加与任务相关的语义，以方便对多媒体信息内容的检索。

研究内容：如何使系统直接从各种媒体中获取信息线索，并将这些线索用于数据库中的检索操作，帮助用户从数据库中检索出合适的多媒体信息对象。

71

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.2 基于内容检索的特点

所谓基于内容检索，就是从媒体数据中提取出特定的信息线索，然后根据这些线索从大量存储在数据库中的媒体中进行查找，检索出具有相似特征的媒体数据出来。

基于内容的检索不同于图像处理、模式识别、计算机视觉、图像理解、语音识别和合成等，但要以这些技术作为其重要的基础。

72

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.2 基于内容检索的特点

基于内容检索的特点：

(1) 基于内容的检索属于多媒体的集合综合技术，不仅要确定是否能够找到，而且还要确定相应的输入输出方法、存储方法、媒体间的组织方法等。它利用认知科学、用户模型、图像处理、模式识别、知识处理、计算机图形学、数据库管理以及信息检索等方法，建立新的媒体数据表示方法和数据模型，采用有效和可靠的查询处理算法，使得用户可以在智能化的查询接口的辅助下完成查询检索工作；

73

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.2 基于内容检索的特点

基于内容检索的特点：

(2) 基于内容检索与模式识别、图像理解、计算机视觉等等学科的重要区别是：它是一种信息检索技术，要以用户可以接受的响应时间、尽量以领域无关的方式检索到所要求的数据。它可以不需要去理解媒体中的对象，更关注的是基于内容的快速查找和发现；

74

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.2 基于内容检索的特点

基于内容检索的特点：

(3) 须从媒体内容中提取信息线索：力图突破传统的基于关键字检索的局限，直接对图像、视频、音频进行分析、抽取特征，使得检索更加接近媒体对象。

(4) 提取特征的方法多种多样：以图像的特征为例，可以提取形状特征，颜色特征，纹理特征，轮廓特征等；

75

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.2 基于内容检索的特点

基于内容检索的特点：

(5) 检索过程人机交互进行：一般来说，人对于特征比较敏感，能迅速分辨出目标的轮廓、音乐的旋律等。但是，对于大量的对象，一方面难以记住这些特征，另一方面人工地从大量的数据中查找目标效率较低。

(6) 基于内容的检索是一种近似的匹配：一般来说，在检索的过程中，采用逐步求精的办法，每一层的中间结果是一个集合，不断减少集合的范围，知道定位目标，这一点与数据库检索的精确匹配算法有明显的不同。

76

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.3 基于内容检索系统的一般结构

基于内容检索结构一般用于多媒体数据库系统之中，也可以单独建立应用系统。从基于内容检索的角度出发，系统由几个大的模块组成：

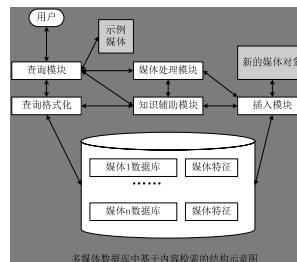
- 组织媒体输入的插入子系统
- 对媒体作特征提取的媒体处理子系统
- 储存插入时获得的特征和相应媒体数据的数据库
- 支持对该媒体的查询子系统

77

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.3 基于内容检索系统的一般结构



78

第八章 多媒体数据库与基于内容的检索

- 4. 基于内容检索系统的结构和方法
- 4.3 基于内容检索系统的一般结构
- 4.3.1 插入子系统

该子系统负责将媒体输入到系统之中，同时根据需要提供用户一种工具，以全自动或半自动（即需要用户干预）的方式对媒体进行分割或节段化，标识出需要的对象或内容关键点，以便有针对性的对目标进行特征提取。

79

第八章 多媒体数据库与基于内容的检索

- 4. 基于内容检索系统的结构和方法
- 4.3 基于内容检索系统的一般结构
- 4.3.2 特征提取子系统

对用户或系统标明的媒体对象进行特征提取处理。

特征提取可以由人完成，例如给出一些描述特征的关键字；也可以通过对应的媒体处理例程完成，提取一些所关心的媒体特征。

提取的特征可以是全局性的，例如整副图像或视频镜头的颜色分布，也可以针对某个内部的对象，如图像中的子区域、视频中的运动对象等。

在提取特征时，往往需要知识护理模块的辅助，由知识库提供有关的领域知识。

80

第八章 多媒体数据库与基于内容的检索

- 4. 基于内容检索系统的结构和方法
- 4.3 基于内容检索系统的一般结构
- 4.3.3 数据库

媒体数据和插入时得到的特征数据分别存入媒体数据库和特征数据库。

媒体库包括各种媒体数据，如图像、视频、音频、文本等。特征库包含这种媒体用户输入的特征和预处理自动提取的特征。

数据库通过组织与媒体相匹配的索引来达到快速搜索的目的，从而可以应用到大规模多媒体数据检索过程中。

81

第八章 多媒体数据库与基于内容的检索

- 4. 基于内容检索系统的结构和方法
- 4.3 基于内容检索系统的一般结构
- 4.3.4 查询子系统

主要以示例查询的方式相用户提供检索接口。

检索允许针对全局对象进行如整副图像、视频镜头等，也允许针对其中的子对象以及任意的组合形式来进行，并将检索返回的结果按相似程度进行排列，如有必要可以进一步查询。

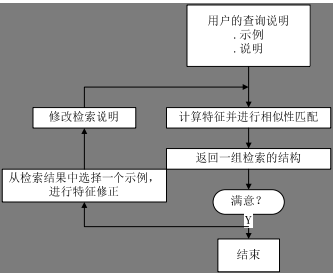
检索主要是相似性检索，模仿人类的认知过程，可以从特征库中寻找匹配的特征，也可以临时计算对象的特征。对于不同的媒体数据类型，具有各自不同的相似性测算法，检索系统中包括一个较为有效可靠的相似度函数集。

82

第八章 多媒体数据库与基于内容的检索

- 4. 基于内容检索系统的结构和方法
- 4.4 基于内容检索的过程和指标
- 4.4.1 检索过程

基于内容检索：
一个逐步求精的过程。



83

第八章 多媒体数据库与基于内容的检索

- 4. 基于内容检索系统的结构和方法
- 4.4 基于内容检索的过程和指标
- 4.4.1 检索过程

初始检索说明：用户开始检索时，要形成一个检索的格式，最初可以用示例的方式给出或者是用特定的查询语言来形成。系统对示例的特征进行提取，或是把用户描述的特征映射为对应的查询参数。

相似性匹配：将特征与特征库中的特征按照一定的匹配算法进行匹配。满足一定相似性的一组候选结果按相似度大小排列返回给用户。

特征调整：用户对系统返回的一组满足初始特征的检索结果进行浏览，挑选出满意的结果，检索过程完成。或者从候选结果中选择一个最接近的示例，进行特征调整，然后形成一个新的查询。
重新检索：逐步缩小查询范围，重新开始。该过程直到用户放弃或得到满意的查询结果时为止。

84

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.4 基于内容检索的过程和指标

4.4.2 特征匹配：过程及方法

特征匹配是基于内容检索最关键的部分。因为媒体的内容语义无法十分精确，所以要采用相似性的匹配方法。

例：

1) 模糊值：二值逻辑在某些情况下不足以表达多媒体对象的特征

2) 特征值分类和计算：

- 用户查询时描述的特征值；
- 从图像数据中直接导出的特征，插入时由系统自动计算提取；
- 需要动态计算的特征值；

85

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.4 基于内容检索的过程和指标

4.4.2 特征匹配：过程及方法

3) 特征调整：初始查询完成后，用户遍历挑选。或从中选出一副图像，针对该图像进行特征调整。

4) 相似特征偏差：相似性匹配必须给出每个特征值可以偏差的程度

5) 相似度及最终求值：“相似度”即相似性的度量结果

86

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.4 基于内容检索的过程和指标

4.4.3 主要指标

基于内容检索由于采用相似性匹配，检索到的对象往往存在一定的误差，这个误差可以用“查到率”和“查准率”来表示。

查到率是指数据库中所有的相关对象是否都查到了。

查准率是指查到的对象是否都是正确的。

87

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.5 媒体的内容语义

图像：图像媒体的常用检索内容主要包括以下一些：

- 颜色：图像颜色的分布、颜色的相互关系、颜色的组成等
- 纹理：图像的纹理结果、方向、组合及对称关系等
- 轮廓：图像轮廓的组成、形状、大小等
- 对象：图像中子对象的关系、数量、属性、旋转等
- 领域内容：某一个领域下的图像内容，例如头像中嘴与眼的相对位置等。

88

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.5 媒体的内容语义

视频：视频可以建立在图像的基础之上，先有图像的内容才可以得到视频的内容。常用的视频内容检索主要有以下一些：

- 镜头：镜头是视频的基本单位，
- 运动对象：查找视频序列中的某一运动对象
- 场景：相同的镜头的寻找和组合等

89

第八章 多媒体数据库与基于内容的检索

4. 基于内容检索系统的结构和方法

4.5 媒体的内容语义

声音：声音的内容检索包括特定模式的查找，特定词、短语、音乐旋律和特定声音的查找等。

例如语音内容的识别，非语音信号的查找等。

90

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.1 基于颜色直方图检索

5.1.1 颜色直方图原理及性质

假设一幅图像G的颜色（或灰度）由N级组成，每一种颜色值用 q_i 记录，具有该颜色值的像素的个数为 h_i ，得到直方图。直方图描述了一幅图像的颜色特征。用直方图描述图像的颜色特征有如下性质：

（1）直方图中的值是统计而来的，描述了该图像关于颜色的数量特征，可以反映图像的部分内容。例，如果是一幅“蓝色的海洋”，蓝色将是像素的主要成分。

（2）直方图丢失了颜色的位置特征。因此，不同的图像可能具有相同的颜色分布，从而也就具有相同的颜色直方图。

91

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.1 基于颜色直方图检索

5.1.1 颜色直方图原理及性质

（3）如果将图像划分成若干子区域，所有子区域的直方图之和等于全图直方图。

（4）很多情况下，由于图像上的背景和前景物体颜色分布会有较大的差异，从而在直方图上会出现双峰特征，但若前景和背景较为接近的图像则不具备该特征。

92

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.1 基于颜色直方图检索

5.1.2 利用颜色直方图进行检索

给出示例图像，得到它的颜色直方图，用该颜色直方图与数据库中的图像颜色直方图进行匹配，最后确定所要找的图像集合。

示例可以用以下的方法给出：

（1）指名颜色的组成，例如。查询“红色30%，蓝色50%”。

（2）指明一副图像：确定示例图像，得到它的颜色直方图，然后用其与数据库中的数据进行匹配。一般来说进行的是相似性匹配。

（3）指明图像中的一个子图：一个子块，一个区域，也可以是利用对象轮廓方法确定的一个对象。利用这个子图确定相应的颜色直方图，再从图像数据库中寻找出类似特征的图。

93

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.1 基于颜色直方图检索

5.1.2 利用颜色直方图进行检索

利用颜色直方图进行查询必须首先确定颜色的级数。当颜色级数很大时，例如24位真彩色，对每一种颜色都进行计算和处理显然不是很合适。一个可行的方法就是减少颜色的数量，例如转换为256色。在匹配的时候也应该采用模糊匹配。

由于颜色直方图不包含位置信息，所以可能会有两幅完全不同的图像具有完全一样的颜色直方图的情况。为此，可将图像进行分割，使其形成若干的子块。这些子块的颜色直方图之和就是整副图像的直方图。子块的数量就可以用来提供某种位置信息。

94

$$\sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.1 基于颜色直方图检索

5.1.3 颜色直方图的相似性匹配

示例图像的颜色直方图 $G(g_1, \dots, g_N)$ 。

数据库中目标图像直方图： $S(S_1, \dots, S_N)$

这两个直方图是否相似，可以用差值表示。例如，借助距离的表示方法：

$$E_d(G, S) = \sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

95

$$\sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.1 基于颜色直方图检索

5.1.3 颜色直方图的相似性匹配

将其规范并简化，1为完全相似，0为完全不相似来确定两个图像的相似性。则相似性可以用下式来描述：

$$\text{Sim}(G, S) = \frac{1}{N} \sum_{i=1}^N \left[1 - \frac{|g_i - S_i|}{\text{Max}(g_i, S_i)} \right]$$

其中，N为颜色级数， g_i, S_i 都大于0。

如果Sim的值靠近1，则相似，否则不相似。

96

$$\sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.1 基于颜色直方图检索

5.1.3 颜色直方图的相似性匹配

直方图中，可能有许多颜色的统计值很小，包括一些不重要的信息。为了消除这种影响，一般来说采用一个阈值加以比较，对不超过这个阈值的颜色不计算到直方图内。如果对其中的某些颜色有重要程度的区别，则可以使用权重因子 W_i 来描述，这样上式就变为：

$$\text{Sim}(G, S) = \frac{1}{N} \sum_{i=1}^N \left\{ W_i \left[1 - \frac{|g_i - S_i|}{\text{Max}(g_i, S_i)} \right] \right\}$$

因为 $0 \leq W_i \leq 1$ ，因此，会引起Sim值的下降。不能很好的反映相似性。因此，可以再做一个调整，从N个颜色值中选取L个最大的单元值进行求和平均。

97

$$\sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.1 基于颜色直方图检索

5.1.3 颜色直方图的相似性匹配

全局直方图除了具有计算简单的特性外，同时还具有对平移和旋转不敏感的优点。但是，全局直方图无法捕捉颜色组成之间的空间关系；事实上，颜色的不同空间分布极大地影响了人们对图像的相似性判断。

基于此，许多研究人员着手研究颜色的空间索引技术。颜色的空间索引技术主要可以分为两大类：

(1) 基于图像空间的固定划分方法，即人为地将图像划分成适当的分块，然后为每个分块提取相应的局部颜色特征。

(2) 基于像素颜色的空间相关性的聚类方法，其出发点是基于如下推断：不同的图像如果在空间相近的位置同时存在颜色相似的大片区域，则这两幅图像具有较大的相似度。

98

$$\sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.2 图像轮廓、纹理特征检索简述

轮廓是图像目标的主要特征。基于轮廓或基于骨架的检索能使用户通过勾画图像的大致轮廓，从数据库中检索出轮廓相似的图像。

取图像的轮廓线：是一个较为困难的工作。一般，采用图像分割、识别目标的前景、背景来得到比较精确的轮廓。

用户在提交检索示例的时候可以自己勾画一个轮廓。

99

$$\sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

第八章 多媒体数据库与基于内容的检索

5. 图像内容分析及检索方法简介

5.2 图像轮廓、纹理特征检索简述

基于纹理的检索：纹理也是图像中重要而又难以描述的特征。很多图像在局部区域内呈现不规则特性，但在整体上表现出规律性，习惯上，把图像的这种局部的不规则而宏观又规律的特性称为纹理。纹理特征包括粗糙性、方向性和对比度。

100

$$\sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

第八章 多媒体数据库与基于内容的检索

6. 视频检索与索引

6.1 视频媒体基本特性

6.1.1 视频序列

视频数据：连续的图像序列。

为了对视频序列进行分类和检索，必须对视频序列的数据结构有所了解。可以这样划分：

镜头 (shot)：包含一个事件或一组连续的动作。每个镜头中的内容都是发生在场景中的。

场景 (scene)：场景“是一组镜头的集合，在内容上包含相似的对象或包含类似的背景。

101

$$\sqrt{\sum_{i=1}^N (g_i - S_i)^2}$$

第八章 多媒体数据库与基于内容的检索

6. 视频检索与索引

6.1 视频媒体基本特性

6.1.2 镜头的切换

镜头的切换：切换点是两个不同镜头之间的分割和衔接。

突变切换或直接切换：没有过渡，直接切换

渐变切换：类似于在PowerPoint中页面的切换，有淡入、淡出、慢转换、擦除转换等。

镜头的运动：推拉、摇、跟踪等。

102

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

6. 视频检索与索引

6.2 镜头检测方法

检测出镜头的分割点。
突变切换：帧与帧之间的直方图差别会比较明显。

渐变类的镜头切换：直方图差值可能不明显。可以考虑采用双重比较法：在一个较大的范围内进行比较；设定两个阈值，一个较高，一个较低。首先用一个较低的阈值来确定出潜在渐变切换序列的起始帧。一旦确定了这个帧，就将其与后续的帧进行比较，逐一和后续的每一帧比较，如果差值是“单调”变化的，则比较进行下去，直到这个单调的过程停止为止。然后，用第二个较高的阈值与这个最后的差值进行比较，如果超过这个阈值，则认为这个不断差值单调增的序列对应一个渐变切换。

103

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

6. 视频检索与索引

6.3 视频索引

建立视频索引的目的：

在视频序列中，对镜头的识别可以把视频分割成一个一个以镜头为基本单位的组合。但是，对镜头的分割需要花费大量的时间。如果要对视频的检索，在存储视频时，应在存放时事先对视频进行分析，找到镜头的切换点，并采用恰当的方法对这些镜头进行表示和安排，以便于基于内容的检索。

另外，由于视频片段中镜头数量会很多，应该对这些镜头进行索引，以便于用户的使用。

104

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

6. 视频检索与索引

6.3 视频索引

视频镜头的表示：

对每个镜头加上文字或符号的标识。但这与图像一样，文字和符号并不能完整的表示镜头的内容，而且在识别镜头时必须人工介入。为较好的表示镜头，可以采用代表帧的方法。

代表帧：描述该镜头的关键图像。

105

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

7. 基于内容的音频检索

7.1 概述

但从整体看，音频内容分为三个级别：最低层的物理样本级、中间层的声学特征级和最高层的语义级，如下图所示。从低级到高级，其内容逐级抽象，内容的表示逐级概括。

106

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

7. 基于内容的音频检索

7.1 概述

在物理样本级，音频内容呈现的是流媒体形式，用户可以通过时间刻度，检索或调用音频的样本数据。如现在常见的音频播放程序接口。

中间层是声学特征级。声学特征是从音频数据中自动抽取的。一些听觉特征表达用户对音频的感知，可以直接用于检索；一些特征用于语音的识别或检测，支持更高层的内容表示。另外还有音频的时空结构。

最高层是语义级，是音频内容、音频对象的概念级描述。具体来说，在这个级别上，音频的内容是语音识别、检测、辨别的结果，音乐旋律和叙事的说明，以及音频对象和概念的描述。

107

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

7. 基于内容的音频检索

7.2 基于语音技术的检索

语音检索是以语音为中心的检索，采用语音识别等处理技术。

- (1)利用大词汇语音识别技术进行检索
- (2)基于子词单元进行检索
- (3)基于识别关键词进行检索
- (4)基于说话人的辨认进行分割

108

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

7. 基于内容的音频检索

7.3 音频检索

音频检索是以波形声音为对象的检索。

(1)声音训练和分类：通过训练来形成一个声音类。用户选择一些表达某类特性的声音例子(样本)，如“脚步声”。

(2)听觉检索：听觉感知特性，如基音和音高等，可以自动提取并用于听觉感知的检索，也可以提取其他能够区分不同声音的声学特征，形成特征矢量用于查询。

(3)音频分割：以上方法适合单体声音的情况，如一小段电话铃声、汽车鸣笛声等。但是，一般的情况是一段录音包含许多类型的声音，由多个部分组成。这需要在处理单体声音之前先分割长段的音频录音。

109

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

7. 基于内容的音频检索

7.4 音乐检索

音乐检索是以音乐为中心的检索，利用音乐的音符和旋律等音乐特性来检索。

110

$$\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}$$

第八章 多媒体数据库与基于内容的检索

小结

从传统数据库到多媒体数据库：需要从体系结构、数据模型、用户接口等方面进行考虑。

基于内容的检索：对内容进行识别和匹配。掌握图像直方图的概念和用于图像相似性比较的方法。

111