

基于内容的视频查询系统研究

蒋薇 尔桂花 戴琼海

(清华大学自动化系 宽带网数字媒体实验室, 北京 100084)

摘要 由于多媒体数据库管理和检索的效率直接决定了人们利用多媒体数据信息的效率,因此随着 MPEG-7 标准的提出,基于内容的图象/视频存储和检索已成为研究的热点。为了快速地对视频进行浏览和检索,在研究基于内容的视频数据库管理和检索等热点问题的基础上,首先使用 MPEG-7 视觉内容描述子和语义描述子来构建视频数据库的语义结构,并结合底层视觉特征和高层语义特征,采用相关反馈机制和半自动权重更新体制来对视频数据库进行管理和检索;然后采用语法分析器来支持自然语言查询;最后在此基础上实现了基于内容的视频数据库的管理和查询系统。实验证明,该系统能够有效地对视频数据进行管理和检索,并且具有一定的智能性和适应性。

关键词 计算机信息管理系统(520·6070) MPEG-7 基于内容的视频描述 语法分析器 相关反馈 权重更新
中图分类号: TP391.3 TN948.61 **文献标识码**: A **文章编号**: 1006-8961(2003)11-1320-07

A Research on MPEG-7 Describing Content-based Video Retrieval System

JIANG Wei, Er Gui-hua, DAI Qiong-hai

(The Department of Automation, Broad Band Network & Digital Multimedia Lab, Tsinghua University, Beijing 100084)

Abstract The management and retrieval of Image/Video database based on their content has become an important search area in recent years. MPEG-7 standard provides a rich set of tools for completely describing multimedia content. In this paper, we explore the issue on content-based video manage and retrieval system. A video databasc described by MPEG-7 video content description and MPEG-7 semantic description is realized. Within this system, natural language comprehension technique is incorporated which makes it possible for user to query conveniently by more natural language, and synonymy lexicon is used to reduce the redundancy of features. We combine Low-level visual feature and high-level semantic description together, and adopt relevance feedback mechanism which enables user to find his target video step by step. Also we use semiautomatic intelligent weight innovation algorithm, which can make the features more and more reasonable. Experiment result shows that we can manage and retrieval video by content expediently and effectively using this system.

Keywords MPEG-7, Content-based video description, Natural language comprehension, Relevance feedback, Semiautomatic weight innovation

0 引言

当今社会,人们已经能够获得大量的多媒体数据,但从这些多媒体数据中获得信息量的大小直接取决于对多媒体数据库进行管理和检索的效率。随

着 MPEG-7 标准的提出,基于内容的图象/视频存储和检索已成为研究的热点。MPEG-7 提供了迄今为止最丰富的一套音视频特征描述,从而使人们对视频内容有了全新的把握^[1~7]。目前已经有很多采用图象/视频的底层特征进行检索的系统^[8~10],但是用户还常常需要根据一些关于视频段的更人性化

基金项目:国家自然科学基金资助项目(60172026);清华大学基础研究基金资助项目(JC2001028);

清华大学深圳研究生院宽带网络中心重点实验室资助项目

收稿日期:2002-11-26; 改回日期:2003-06-26

的特征进行检索,如“高兴”、“生气”等,由于这些特征属于视频段的高层语义特征,因此计算机日前尚无法自动提取. 尽管现在这些高层语义特征只能用关键字标注来表示,而且只能由人机交互进行半自动标注,但是这些特征却包含了人类对图象内容的理解,并含有大量的信息,由于使用关键字进行辅助查询,能够提高检索速度、准确度以及检索效率^[11~13],因此如何有效地利用这类关键字特征来对图象/视频数据库进行检索是很值得研究的问题.

本文主要针对视频数据库中这类高层语义特征的标注和检索进行了研究,并使用 MPEG-7 内容描述子来构建视频数据库的语义结构,因而具有可扩展性. 由于传统的关系数据库文本检索方式是在检索的时候使用关系运算符“and”、“or”等把要查询的关键字输入给系统,而不能支持更自然的语言查询,为此本文引入了语法分析器,以便能够支持自然语言查询,同时由于该系统将视频的底层视觉特征描述和高层语义信息描述相结合来提供相关反馈的查询机制,并且采用半自动更新特征权重的机制来在相关反馈的同时更新特征权重,以自动剔除不合理的特征和加重合理的特征,因此具有一定的智能性和适应性.

1 MPEG-7 简介

MPEG-7 标准,又叫“多媒体内容描述接口”,其提供了迄今为止最丰富的一套音视频特征描述. 它不仅提供了通用的、灵活的、可扩展的多媒体内容描述框架,而且标准化了一个描述符 D(Descriptor)集合、一个描述方案 DS(Description Scheme)集合、一种描述定义语言 DDL(Description Definition Language),以及对描述进行编码的多种方法和工具. 其中描述符表示了多媒体对象的某种特征;描述方案规定了在某个描述中可用的描述符,描述符的生成,以及描述符与描述符、描述符与描述方案、描述方案与描述方案之间的关系的结构和语义;描述定义语言是用来生成新的描述方案或者描述符的一种语言,它也可以扩展或者修改已有的描述方案. 它们之间的关系如图 1 所示.

MPEG-7 的研究范围仅限于多媒体对象描述的生成,而具体采用什么样的特征提取算法来提取描述,以及如何使用搜索引擎实现搜索则都不在 MPEG-7 标准的范围内(如图 2 所示).

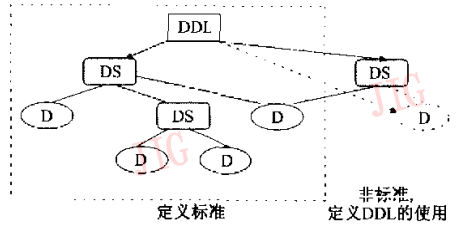


图 1 描述方案和描述符及描述定义语言关系图

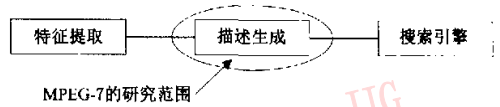


图 2 MPEG-7 标准的范围图

2 MPEG 视频语义描述

本系统使用了基于 MPEG-7 内容描述的概念来构建视频数据库的语义结构,下面对 MPEG-7 的视频内容描述作简单介绍. MPEG-7 视频描述有许多方面的内容,这里仅介绍系统使用的部分,更详细的描述可以参考文献^[2,4~7].

2.1 视频结构和语义描述

为构建视频段的语义描述体系,本系统使用了视频段描述(Video Segment DS)和语义描述(Video Semantic DS),它们通过连接相关联,以使得视频内容可以同时基于结构和语义来描述.

(1) Video Segment DS

它代表了一段视频单元,也是一个抽象类型,它可以理解为一组视频帧. Segment DS 可以组织成层次结构,并可生成一个表,用来对视频内容进行标注索引(见图 3). Segment DS 可以使用文本注解来进行进一步的语义描述. 本系统选用内容实体来描述类型(Content Entity Description Type)中的视频类型,其具体定义如下:

Segment DS:

$$V = \{V-ID, SV, Annotation\};$$

$$SV = \{SV-ID, SV, Annotation\} \text{ 或}$$

$$SV = \{SV-ID, Sc, Annotation\}$$

其中, V 表示视频段, SV 表示子视频段, Sc 表示场景. 用文本注释(Annotation)来记录上面描述的由用户自己定义的语义特征.

(2) Video Semantic DS

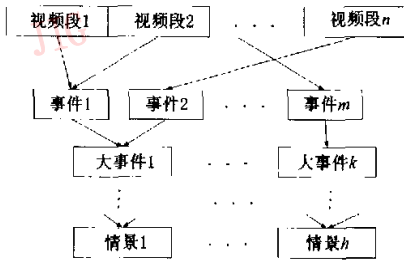


图3 系统使用 Segment DS 描述的视频结构

它包括各种语义实体,如场景(Scene)、对象(Object),对象代理(AgentObject)、事件(Event)、地点(Locality)、时间(Time)、运动(Action)、状态(State)和语义(Semantic)等。Semantic 用于描述视频隐含的语义信息;Object 用于描述可感知的或抽象的实体;AgentObject 用于描述一个人或组织,或者有个性的物体;Event 用于描述叙述世界的某时间某地点发生的事情;State 用于描述某个语义实体的状态;Locality 和 Time 则分别用于描述叙述世界的地点和时间(见图4)。具体定义如下:

Semantic DS:

Sc = {Sc-ID, E, AO, O, T, L};

E = {E-ID, O, AO, Se, T, L, SS};

O = {O-ID, A, T, L, SS}

AO = {AO-ID, A, T, L, SS}

Se = {Sc-ID, Sentence};

L = {L-ID, Annotation};

A = {A-ID, Annotation};

T = {T-ID, Annotation};

SS = {SS-ID, Annotation}

其中,Sc 表示情节,E 表示事件,O 表示对象,AO 表

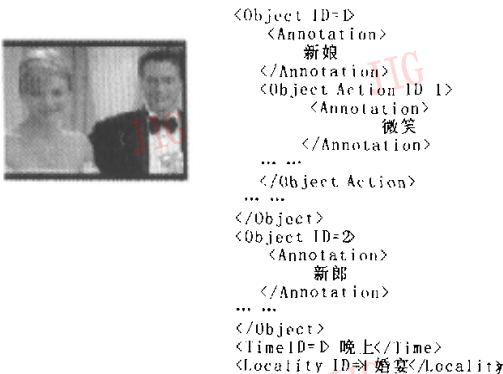


图4 系统使用 Semantic DS 描述的视频结构

示对象代理,Se 表示语义,T 表示时间 L 表示地点,SS 表示状态,A 表示运动.

2.2 底层视觉特征描述

本系统使用了 MPEG-7 的视觉颜色描述 (Visual Color DS)和视觉纹理描述 (Visual Texture DS)来共同描述视频段的底层视觉信息.

(1) Visual Color DS

系统是使用颜色直方图(Global Histogram DS, 256 维),颜色矩(Color Moment DS,3 维),颜色布局 (Color Layout DS, 192 维),和颜色结构 (Color Structure DS,184 维),组成 635 维颜色特征描述.

(2) Visual Texture DS

系统是使用边缘直方图 (Edge Histogram DS, 80 维)组成纹理特征描述.

最后,系统将颜色特征和纹理特征共同组成 715 维的底层视觉特征,并对视频段的关键帧进行底层视觉特征描述,用以代表整个视频段的底层视觉特征.

3 视频管理查询系统

在使用 MPEG-7 内容描述的概念来构建视频数据库语义结构的基础上,实现了视频管理查询系统.该系统的整体结构如图5所示.系统分为建库部分和查询部分:其中建库部分使用上面的 MPEG-7 内容描述概念来构建数据库,并建立视频索引;而查询部分则结合了语法分析器、相关反馈机制和动态更新权重机制,以便使用户能够方便有效地进行视频管理和查询.

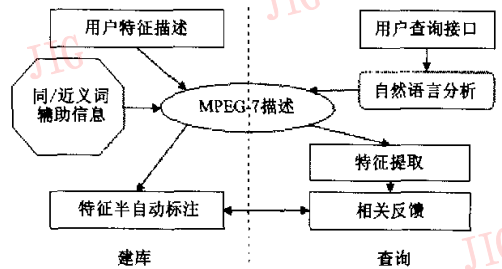


图5 系统结构图

3.1 语法分析器

自然语言理解分为词法分析、句法分析和语义分析3个层次.其中句法分析是对句子和短语的结构进行分析^[14],而人工智能领域则将句法的形式化

定义为: $G=(K, \Sigma, I, C)$. 其中, K 为有穷非空集, Σ 称作总词汇表; Σ 为 V 的一个非空子集, 称作终结字母表; $U=K-\Sigma$ 称作非终结字母表; I 为有穷产生式集 $\alpha \rightarrow \beta, \alpha \in KNK, \beta \in K, \sim$ 表示它前面的字符可以重复出现任意次; C 为非终结字母表的一个元素. 句法分析树的例子如图 6 所示. 为了支持用户的自然语言查询, 可在系统中使用以上的句法分析树来处理用户的查询语句, 并用回溯算法分析句子成分, 以提取关键信息(见图 7).

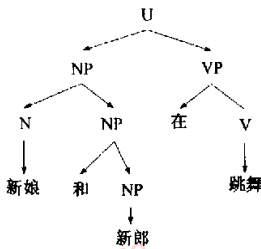


图 6 句法分析树示例

U 为非终结字母表, NP 为名词短语, VP 为动词短语, N 为名词, V 为动词

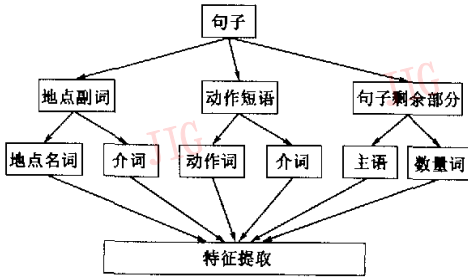


图 7 系统句法分析树结构

从查询语句中提取主语、谓语、宾语以及定语、地点状语和部分补语

3.2 相关反馈机制

相关反馈机制在基于内容的图象/视频检索中是一种十分有效的人机交互的渐进检索方式^[15,16]. Rocchio 在文献^[17]中给出了给予底层视觉特征相关反馈的查询点更新公式

$$B_{new} = \alpha B + \beta \left(\frac{1}{N_{image}^{(R)}} \sum_{i \in D_{image}^{(R)}} D_i \right) - \gamma \left(\frac{1}{N_{image}^{(IR)}} \sum_{i \in D_{image}^{(IR)}} D_i \right) \quad (1)$$

其中, $D_{image}^{(R)}$ 和 $D_{image}^{(IR)}$ 分别表示正例和反例的图象标号集合, $N_{image}^{(R)}$ 和 $N_{image}^{(IR)}$ 是 $D_{image}^{(R)}$ 和 $D_{image}^{(IR)}$ 中的图象数目; 在底层的视觉特征表示中, 每幅图象都被看作高维

空间中的一个点, 并用一个特征向量表示, D_i 是第 i 幅图象的特征向量, B 是查询目标(查询点)在该高维空间中的表示. α, β, γ 是权重, 表示标注示例和关键字的不同重要性, 且 $\alpha + \beta + \gamma = 1$. 当 $\alpha = 0$, 为完全基于示例的查询; 当 $\beta = \gamma = 0$, 为完全基于关键字的查询.

在式(1)的基础上, 针对所研究的高层语义检索问题, 参考文献^[8, 16], 将公式(1)修改如下

$$Q_j = \alpha d_j + \beta \left[\frac{1}{N_{video}^{(R)}} \sum_{i \in D_{video}^{(R)}} d_{i,j}^{(R)} \right] - \gamma \left[\frac{1}{N_{video}^{(IR)}} \sum_{i \in D_{video}^{(IR)}} d_{i,j}^{(IR)} \right] \quad (2)$$

其中, $D_{video}^{(R)}$ 和 $D_{video}^{(IR)}$ 分别表示正例和反例的视频段标号集合, $N_{video}^{(R)}$ 和 $N_{video}^{(IR)}$ 是 $D_{video}^{(R)}$ 和 $D_{video}^{(IR)}$ 的视频段数目, 用 x 表示视频段元数据, Q_j 是数据库中第 j 段视频段 x_j 与当前查找对象之间的距离, 即 x_j 符合当前查找要求的程度(越小的 Q_j , 越符合当前的查找要求); d_j 表示所查找关键字与视频段 x_j 的距离, $d_j = 1 - P(K_i/M_j)$, 其中, $P(K_i/M_j)$ 为 M_j 中具有属性 K_i 的概率; $d_{i,j}^{(R)}$ 和 $d_{i,j}^{(IR)}$ 分别是视频段 x_j 和正例或反例 x_i 的距离

$$d_{i,j}^{(R)} = \frac{\sum_{n \in A_{i,j}^{(R)}} W_j(n)}{\sum_{n \in A^{(R)}} W_j(n)}, \quad d_{i,j}^{(IR)} = \frac{\sum_{n \in A_{i,j}^{(IR)}} W_j(n)}{\sum_{n \in A^{(IR)}} W_j(n)} \quad (3)$$

其中, $A_{i,j}^{(R)}$ 和 $A_{i,j}^{(IR)}$ 分别为视频段 x_j 中与正例或反例 x_i 具有不同特征的标号集合, $A^{(R)}$ 和 $A^{(IR)}$ 分别为正例或反例 x_i 中所有与视频段 x_j 具有不同特征的标号集合, $W_j(n)$ 表示视频段 x_j 中属性 K_i 的相对权重.

这样利用式(2)就可连续进行相关反馈. 另外, 根据用户输入的要查找的属性计算 d_j , 同时根据用户反馈的正反例集合即可由式(2)计算相似距离 Q_j , 并把距离小于一定值的视频段返回, 若用户不满意, 则继续反馈, 以便逐渐接近用户查找的真正目标.

3.3 半自动权重更新体制

由于用户标注的特征可能是不准确的, 因此可能并不能反映视频的真正属性. Zhu 等在文献^[18]中提出了半自动更新权重的机制. 该文假设用户查询语句中的第 i 个属性 K_i , 在数据库中共有 M 段视频具有该属性, 其中第 j 个视频段 x_j 中共有 m 个属性; $P(K_i|x_j)$ 表示视频段 x_j 中具有 K_i 属性的概率,

$W_j(n)$ 为视频段 x_j 中属性 K_i 的相对权重, 所以有

$$P(K_i/x_j) = \frac{W_j(n)}{\sum_{i=1}^m W_j(n)} \quad (4)$$

由式(4)可见, 概率越大的属性越能表现视频的真正内容. 当用户输入某属性, 并把一些视频作为正例或反例时, 正例具有该属性的权重增加, 反例具有该属性的权重减少, 并将权重小于一定值的属性与视频段的连接删除, 这就是权重的半自动更新过程. 由于这种方法能自动剔除不合理的特征, 并能加重合理的特征, 因而具有自我适应能力.

实际应用中, 从特征的最初标注到权重的自动更新都存在一个问题, 就是同义词/近义词可能带来的特征冗余. 针对这个问题, 在系统中采用了同义词/近义词词表, 使同义的或近义的关键字算是同一个特征, 这样就减少了同义词/近义词的冗余干扰.

3.4 底层视觉特征的检索

根据底层特征的检索使用文献[16]中提出的对象模型表示. 每个对象(视频关键帧)都表示为 $O = \{x, F, R\}$ 模型, x 是图象元数据; $F = \{f_i\} (i = 1, \dots, k)$, f_i 表示图象的第 i 种底层特征, 如颜色和纹理特征, 共有 k 种特征, $R = \{r_i\} (i = 1, 2, \dots, k)$, 其中, r_i 是特征 f_i 的特征代表(representation), 假设 f_i 有 m 种特征代表, 则有 $r_i = (r_{i1}, \dots, r_{i2}, \dots, r_{im})$, 如直方图和颜色矩就是颜色特征的两类代表, 每个特征代表 r_{ij} 本身是一个 n 维的特征向量, 即 $r_{ij} = [r_{ij1}, \dots, r_{ijn}]^T$. 如颜色矩为 9 维特征向量. 令 $O^{(q)} = \{x^{(q)}, F^{(q)}, R^{(q)}\}$ 表示当前的查询示例图象, 其中, $F^{(q)} = \{f_i^{(q)}\}$, $R^{(q)} = \{r_{ij}^{(q)}\}$, $r_{ij}^{(q)} = [r_{ij1}^{(q)}, \dots, r_{ijn}^{(q)}]^T$, 则数据库中每个对象的特征代表和查询对象的相应特征代表之间的相似度为: $S(r_{ij}, r_{ij}^{(q)}) = K_{ij}(r_{ij}, r_{ij}^{(q)}, w_{ij})$, $w_{ij} = [w_{ij1}, \dots, w_{ijn}]^T$ 为一组权重系数. 为方便计算, 系统中设定 $w_{ij,n} = 1/n$, 则可使用欧氏距离来计算相似度 $S(r_{ij}, r_{ij}^{(q)})$, 即

$$S(r_{ij}, r_{ij}^{(q)}) = \frac{1}{n} \sqrt{\sum_{l=1}^n (r_{ij,l} - r_{ij,l}^{(q)})^2} \quad (5)$$

由每个特征代表的相似度加权组合成特征的相似度: $S(f_i, f_i^{(q)}) = \sum_j w_{ij} S(r_{ij}, r_{ij}^{(q)})$, 其中, w_{ij} 为每个特征代表的权重. 这里, 令 w_{ij} 反比于正例视频帧在该维特征上的标准方差, 即

$$w_{ij} = 1/\sigma_{ij} \quad (6)$$

其中, σ_{ij} 是正例视频的第 j 个特征代表在第 i 个特

征上的标准方差, 也就是说, 用户反馈中各正例视频取值不一致的特征(σ_{ij} 大)可认为是不重要的(w_{ij} 小), 而取值一致的特征(σ_{ij} 小)则认为是重要的(w_{ij} 大). 因此有

$$S(f_i, f_i^{(q)}) = \sum_{j=1}^m \frac{1}{\sigma_{ij}} S(r_{ij}, r_{ij}^{(q)}) \quad (7)$$

而整个对象的相似度 S 就是将所有特征的相似度进行加权组合: $S = \sum_i w_i S(f_i, f_i^{(q)})$, w_i 表示每个特征的权重, 系统中简单设定为 $w_i = 1/k$, 于是就得到了如下基于底层特征的相似度量

$$S = \sum_{i=1}^k \frac{1}{k} S(f_i, f_i^{(q)}) \quad (8)$$

系统允许用户同时使用高层语义描述(自然语言)和底层视觉特征(示例)进行检索. 如果用户输入查询语句, 则根据查询语句通过高层语义结构来进行检索, 这样, 用户既可以在高层语义检索的基础上, 通过底层特征的示例来进行查询, 又可以由相关反馈来对查询结果进行微调, 另外, 用户也可以不输入查询语句, 只使用底层特征来进行检索. 虽然底层特征的检索结果并不影响上面所说的高层语义标注的半自动权重更新, 但底层特征的引入却使得系统更能适应不同用户检索的不同需求.

4 系统实现

系统以 Oracle 视频数据库为实验平台, 使用 50 部电影作为实验数据. 系统采用 MPEG-7 视频语义描述的概念来构建数据库的语义结构, 并将用户输入的特征放到语义结构中(如图 8 所示). 另外, 利用自然语言查询的一次反馈实验结果如图 9 所示, 利用底层特征微调检索的结果如图 10 所示. 由图中可

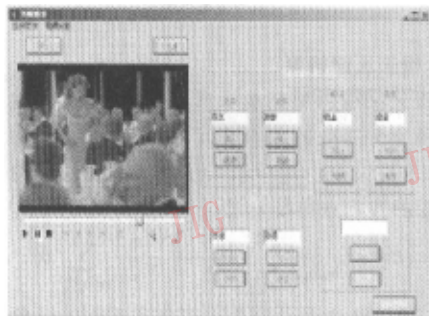


图 8 使用 MPEG-7 内容描述概念构建的系统语义结构



图 9 自然语言查询一次反馈结果
(查询语句“人们在微笑”得到的实验结果)

以看到:(1)系统使用 MPEG-7 描述语言进行描述是为了便于以后扩展使用;(2)使用语法分析器,用户能使用更自然语言更方便的查询;(3)由于采用相关反馈和权重的半自动更新体制,能使用户可以动态地在相关反馈的同时进行语义特征的补充,且其补充特征的合理性可由权重更新体制来管理,因此具有一定的智能性和适应性;(4)采用同/近义词词表来减少特征冗余;(5)若能同时结合高层语义和底层特征进行检索,则更能适应用户的不同需求。

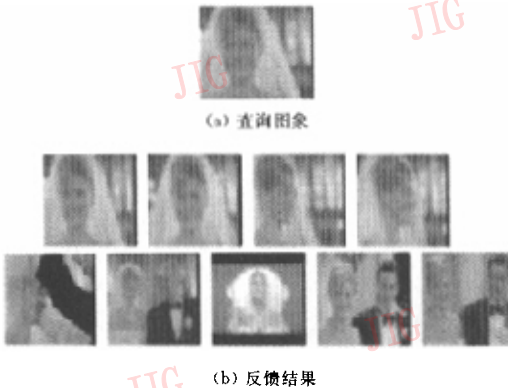


图 10 根据底层特征描述对图(9)结果进行
微调的一次反馈结果
(在图 9 的查询结果基础上,点击(a)图,根据底层视觉特征
得到的一次反馈结果)

5 结 论

本文研究了基于内容的视频管理和查询,同时实现了利用 MPEG-7 视频内容描述子和语义描述

子描述的系统。由于视频的底层特征和高层语义特征的结合仍是基于内容视频数据库研究的热点^[11~13,19],因此将来要在这方面作更进一步研究。另外,关系数据库的数据分割和面向对象数据库的对象分割也还是开放问题^[20],将来也要在此系统上作进一步研究。

参 考 文 献

- Martinez J M. Overview of the MPEG-7 Standard[S]. ISO/JTC1/SC29/WG11/N4509, International Organization for Standardization, Pattaya, Thailand, December 2001.
- MPEG-7 Description Scheme Group. MPEG-7 description schemes [S]. ISO/IEC/JTC1/SC29/WG11/N2844, International Organization for Standardization, Vancouver Canada, July 1999.
- Day Neil, Martinez JM. Introduction to MPEG-7 [S]. ISO/IEC/JTC1/SC29/WG11/N4325, International Organization for Standardization, Sydney, Australia, July 2001.
- Vetro A. MPEG 7 Applications Document [S]. ISO/IEC/JTC1/SC29/WG11/N3934, International Organization for Standardization, Pisa Italy, January 2001.
- MPEG-7 Video Group, MPEG-7 Visual Part of Experimentation Model [S]. ISO/IEC/JTC1/SC29/WG11/N2822, International Organization for Standardization, Vancouver Canada, 1999.
- Cieplinski L, Kim M, Ohm J R. et al. MPEG-7 multimedia content description interface-part 3 Visual[S]. ISO/IEC/JTC1/SC29/WG11/N4062, International Organization for Standardization, Singapore, March 2001.
- Sikora T. The MPEG-7 visual standard for content description—an overview [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001,11(6):696~702.
- Yang Jun, Liu Wenyin, Zhang Hongjiang et al. An approach to semantics-based image retrieval and browsing [A]. In: International Workshop on Multimedia Database Systems[C], Taipei, China, September 2001.
- Gaurav Aggarwal, Ashwin T V, Sugata Ghosal. An image system retrieval with automatic query modification[J]. IEEE Transactions on Multimedia, 2002,4(2):201~214.
- Koskela M, Laaksonen J, Oja E. PicSOM-self-organizing image retrieval with MPEG-7 content descriptors [J]. IEEE Transactions on Neural Networks, 2002,13(4):841~853.
- Zhou Xiangsean, Huang T S. Unifying keywords and visual contents in image retrieval[J]. IEEE Multimedia, 2002,9(2):23~33.
- Shekholeslami G, Chang Wendy, Zhang Aidong. SemQuery: semantic clustering and querying on heterogeneous features for visual data [J]. IEEE Transactions on Knowledge and Data Engineering, 2002,14(5):988~1002.
- Ogle V E, Stonebraker M. Chabot: Retrieval from a relational

- database of images[J]. IEEE Computer, 1995,28(9):40~48.
- 14 蔡自兴,徐光祐. 人工智能及其应用(第二版)[M]. 北京:清华大学出版社,2000.
- 15 Rui Y, Huang T S, Ortega *et al.* Relevance feedback: a powerful tool for interactive content-based image retrieval[J]. IEEE Transactions on Circuits and Video Technology, 1998,8(5):644~655.
- 16 Zhang Hongjiang, Chen Zheng, Liu Wenyin *et al.* Relevance feedback in content-based image search [EB/OL]. http://research.microsoft.com/asia/dload_files/group/mcomputing/nit01hj-4th.pdf, 2001.
- 17 Rocchio J. Relevant feedBack in information retrieval[A]. In: Geraxd Slaton, editor, The smart retrieval system-experiments in automatic document processing[C], Cleffs Englewood. NJ, USA;Prentice-Hall, 1971;313~323.
- 18 Zhu Xingquan, Liu Wenyin, Zhang Hongjiang *et al.* An image retrieval and semi-automatic annotation scheme for large image databases on the web [A]. In: Proceeding of 13th SPIE symposium on Electronic Imaging [C], San Jose, CA, USA, 2001,4311:168-177.
- 19 Shermann S M, Chan Li Qing, Wu Yi *et al.* Accommodating hybrid retrieval in a comprelhensive video database management system[J]. IEEE Transactions on Multimedia, 2002,4(2):146~159.
- 20 Karlapalem K. Li Qing. A framework for class partitioning in object oriented databases [J]. Distributed and Parallel Databases, 2000,8(3):317~350.



蒋薇 1979年出生,2002年获清华大学自动化系工学学士学位,现在清华大学自动化系控制理论与控制工程研究所宽带网数字媒体技术实验室攻读工学硕士学位.主要研究方向为数字图象处理、多媒体数据库和计算机网络技术等.



尔桂花 清华大学自动化系副教授,1984年获天津大学自动化系工学学士学位,1989年获北京理工大学自动控制系工学硕士学位.长期从事自动控制专业教学和科研工作.主要研究方向为自动控制系统CAD、变频调速、现场总线、网络安全、多媒体信息编解码、多媒体数据库等.

戴琼涛 现为清华大学副教授,清华大学自动化系宽带网数字媒体实验室主任.主要研究领域为宽带网络、接入网、复杂流媒体、IP网络QoS及其网络行为学的研究、IPV6的安全机制及协议研究、宽带信息网络业务资源模型及分析.