

# 基于图模型的足球视频语义建模方法

王昊冉 白亮 老松杨

(国防科学技术大学信息系统与管理学院 长沙 410073)

**摘要** 视频低层特征和高层语义之间存在的“语义鸿沟”是视频分析与检索应用研究的一个技术瓶颈问题。通过深入分析,提出了足球视频语义内容的时空关联特性,尝试引入图模型方法来建模这种语义关联,提出了视频语义图(VSG)的语义建模方法和基于 DFS 的视频语义图匹配算法,并分析了算法的复杂性。实验结果显示,该方法能够有效解决视频语义内容建模和分析匹配问题。

**关键词** 视频语义图,图模型,图匹配

**中图分类号** TP391 **文献标识码** A

## Approach Based on Graphics Model for Semantic Modeling in Soccer Video

WANG Hao-ran BAI Liang LAO Song-yang

(School of Information System and Management, National University of Defense Technology, Changsha 410073, China)

**Abstract** The Semantic Gap which exists between video low-level features and high-level semantics is the bottleneck of the video analysis and retrieval. To bridge the gap in football matches videos, the graphical model was brought in, and the Video Semantic Graph was proposed to modeling the high-level semantics. In addition, a DFS-based VSG matching algorithm was proposed, the time complexity and the space complexity were analyzed. The results of the experiment demonstrate the validity of the presented method, which can solve the problems of video semantic modeling and matching.

**Keywords** Video semantic graph, Graphical model, Graph model matching

## 1 引言

伴随信息社会的发展,视频应用的需求在迅速增长。特别是,人们不愿花费很多时间去观看一场完整的体育比赛,更倾向于观看这场比赛中的精彩事件,比如足球比赛中的进球事件。然而,传统视频内容分析方法抽取客观存在的感知特征,而用户所消费的往往是语义内容,这就造成了计算机自动分析与用户需求之间的矛盾。人们把这种矛盾称之为“语义鸿沟”<sup>[1,8]</sup>。针对视频高层语义建模,来跨越这个鸿沟,是视频研究领域的前沿问题。

其中,常常用到隐马尔可夫模型(HMM)<sup>[2]</sup>、支持向量机(SVM)<sup>[3]</sup>、Bayesian 网络<sup>[9]</sup>等数学工具对不同的视频语义内容建立模型。文献[4]提出用本体(Ontology)对视频概念内容建模。本体是一种有效的概念建模和知识表示工具,视频语义内容分析方面,本体应用主要集中在构建特定领域概念本体、辅助视频概念标注(Annotation)和索引(Indexing)。

在计算机研究和应用中,图(Graph)常被用来描述复杂的数据对象结构或者基本数据对象之间的关系。近年来,在视频语义内容建模研究中,部分学者将图模型引入其中,试图找到能够自动化地索引和检索视频的方法。例如, Lee 等人<sup>[3]</sup>使用图来建模视频中的前景和背景对象,他们对得到的

图使用编辑距离度量(edit distance metric)做聚类,再通过多层次的索引完成相似性查询。

本文用图来建模视频语义内容,分析了足球比赛视频的语义内容时空关联特性。通过研究发现,用图模型描述这种时空关联特性是合适的。研究了基于图模型的足球视频语义建模方法,提出了视频语义图(Video Semantic Graph, VSG)的概念,并分析了图匹配的优先算法(DFS),提出了基于图匹配的视频检索算法,在一定程度上构建了跨越“语义鸿沟”的桥梁。

## 2 相关工作

体育视频的编辑特性,决定了其在时空上的相关属性。通过收集大量体育视频片段,本文发现体育视频语义内容的时间相关性,如图 1 所示。空间上的相关性主要是由体育比赛的场地决定的,可产生远景和近景等镜头。

慢镜头是镜头中的一类,是视频中的一类重要语义内容,表现为帧间具有较大的相似性。探测慢镜头对于视频内容的分析具有重要的作用。例如在体育赛事中,慢镜头是导演根据现场情况对比赛中的精彩片断多个不同视角的重放和慢放,可以让观众更加清晰地了解到比赛中一些感兴趣的画面。慢镜头对应着体育比赛中的重要事件,它本身包含强烈的语

到稿日期:2010-07-10 返修日期:2010-11-13 本文受国家自然科学基金青年基金项目(60902094)资助。

王昊冉(1988—),男,硕士生,主要研究方向为指挥控制与指挥决策分析,E-mail:w19881120@163.com;白亮(1978—),男,博士,讲师,主要研究方向为数字音视频处理、多媒体信息系统;老松杨(1968—),男,教授,博士生导师,主要研究方向为多媒体信息系统、指挥决策分析。

义信息,可以作为体育视频中的一个基本语义单元。本文采用在 MPEG 压缩域内进行的基于镜头密度和帧间差的慢镜头探测方法<sup>[10]</sup>。该探测方法的第一步是基于宏块类型信息的镜头探测,在镜头探测的基础上定义和计算镜头密度,探测慢镜头可能出现的位置;第二步,计算压缩域下的连续帧间差,从候选慢镜头中探测慢镜头。本方法与基于帧间差的慢镜头探测方法在探测性能上相当,平均查全率大于 85%,平均准确率大于 90%。但是在探测速度上,本文采用的方法,由于避免了逆 DCT 变换,所需时间仅为基于帧间差的慢镜头探测算法的 1/10,明显快于基于帧间差的方法。

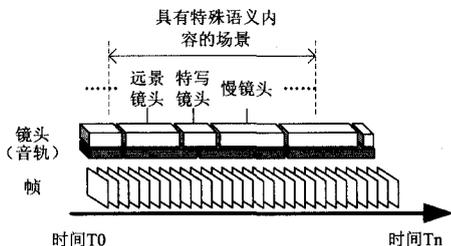


图 1 体育视频语义内容时间相关性

在体育视频伴随音频处理方面,采用文献[4]提出的基于支持向量机的音频概念探测方法,对体育视频中的欢呼声进行重点检测。

提出的基于图模型的视频语义内容建模方法,正是在上述相关工作的基础上展开的,其运用图模型及其理论实现对高层视频语义内容的建模和检索。

### 3 图模型的相关理论

这里,给出图的相关定义。

**定义 1**(图, graph) 所谓图  $G$  是一个三元组,记作  $G = \langle V(G), E(G), \varphi(G) \rangle$ 。其中,

(1)  $V(G) = \{v_1, v_2, \dots, v_n\}, V(G) \neq \emptyset$ , 称为图  $G$  的结点集合(vertex set)。

(2)  $E(G) = \{e_1, e_2, \dots, e_m\}$  是  $G$  的边集合(edge set), 其中  $e_i$  为  $\langle v_j, v_k \rangle$  或  $\langle v_j, v_l \rangle$ 。若  $e_i$  为  $\langle v_i, v_i \rangle$ , 称  $e_i$  为以  $v_j$  和  $v_l$  为端点(end vertices)的无向边(undirected edge); 若  $e_i$  为  $\langle v_j, v_k \rangle$ , 则称  $e_i$  为以  $v_j$  为起点(origin)、 $v_k$  为终点(terminus)的有向边(directed edge)。

(3)  $\varphi(G): E \rightarrow V \times V$  称为关联函数(incidence function)。

**定义 2**(有向图, directed graph) 每一条边都是有向边的图称为有向图。

**定义 3** 设  $D$  是任意有向图,  $x$  为  $G$  的任一结点, 射入  $x$  的边数称为  $x$  的入度(in-degree), 记作  $\deg^+(x)$  或  $d^+(x)$ ; 射出  $x$  的边数称为  $x$  的出度(out-degree), 记作  $\deg^-(x)$  或  $d^-(x)$ 。有向图  $D$  中, 任意结点  $x$  的度数  $\deg(x) = \deg^+(x) + \deg^-(x)$ 。

**定义 4** 设  $G = \langle V, E \rangle$  是一个有  $n$  个结点且没有平行边的有向图, 则一个  $n \times n$  阶的矩阵  $M = [m_{ij}]$  称为图  $G$  的邻接矩阵(adjacent matrix), 且

$$m_{ij} = \begin{cases} 1, & \text{如果有向边 } \langle v_i, v_j \rangle \in E \\ 0, & \text{如果没有从 } v_i \text{ 指向 } v_j \text{ 的边} \end{cases}$$

但是, 一个杂乱无章的图(见图 2), 如何能够表示复杂的视频语义内容呢? 需要解决 3 个问题: 第一, 如何定义每个建模元素的图形符号; 第二, 如何定义图形符号之间的位置关

系; 第三, 如何将表示法映射到抽象语法。

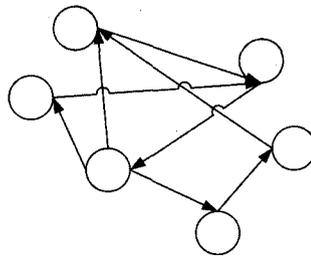


图 2 一个拓扑结构杂乱的图

图形化的视频语义内容建模方法的定义: 构造图  $G = \langle V, E \rangle$  是这样图, 结点  $V = \{v_1, v_2, \dots, v_n\}$  表示各种视频语义内容, 边  $E = \{e_1, e_2, \dots, e_m\}$  表示各个结点之间的关系, 这种构造图  $G$  的方法, 称为图形化的视频语义内容建模。本文中图  $G$  称为视频语义图(Video Semantic Graph, VSG)。

### 4 基于图模型的视频语义建模

不同的体育视频具有相似的语义镜头, 图 3 分别展示了足球视频、排球视频和网球视频中的远景、近景和特写镜头的关键帧。根据提出的视频语义图(VSG)的定义, 针对足球视频语义内容进行分析与建模。

	远景镜头	中景镜头	特写镜头
足球视频			
排球视频			
网球视频			

图 3 不同体育视频中的语义镜头

#### 4.1 足球视频高层语义建模

足球比赛视频具有以下显著特征, 有利于足球视频语义内容分析的建模与实现。

- 足球比赛视频的结构性强;
  - 足球比赛视频的拍摄编辑规则为语义内容分析提供了有利的线索;
  - 足球比赛规则为语义内容分析提供了重要依据。
- 足球视频中常见的事件单元, 如表 1 所列。

表 1 足球视频中常见的事件单元

1	球员特写
2	慢镜头回放
3	声音能量变化
4	比分牌变化
5	开球区镜头
6	罚球区镜头
7	球门镜头
8	远景镜头
9	近景镜头

提出了用结点  $V = \{v_1, v_2, \dots, v_n\}$  表示事件, 边  $E = \{e_1, e_2, \dots, e_m\}$  表示在视频帧序列中的时序关系, 构建足球视频语义图的方法。由表 1 可以清晰地看出足球视频中复杂的语义

事件,但这些事件的出现呈现有规律的周期性。因此,可以针对这些语义事件定义不同的结点,如图4所示。

0	PTV	球员特写
1	SM	慢镜头回放
2	APC	声音能量变化
3	SC	比分牌变化
4	KOA	开球区镜头
5	PKTV	罚球区镜头
6	GV	球门镜头
7	LV	远景镜头
8	CV	近景镜头
9		.....
		语义之间关系(时间关系、空间关系等)

图4 面向足球视频语义内容建模图模型定义

例如,在足球比赛中判断进球这一语义。根据体育视频的编辑特点,进球可以描述为“球门”——“运动员特写”——“音频能量提高”——“慢镜头回放”——“比分牌变化”,如图5所示。根据视频中“进球”语义的时序变化规律,利用本文提出的图模型描述方法,可得到“进球”的视频语义图,如图6所示。

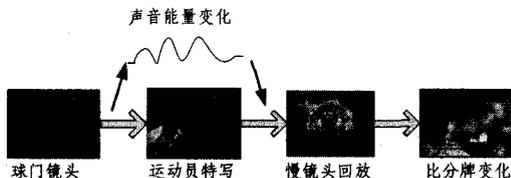


图5 进球语义概念实例

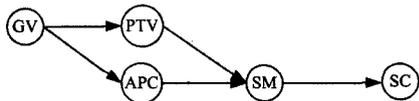


图6 进球的视频语义图

再例如,表示点球这一语义。点球可以描述为“罚球区特写”——“球门镜头”——“音频能量提高”——“比分牌变化”——“运动员特写”——“慢镜头回放”,如图7所示。则点球的视频语义图如图8所示。

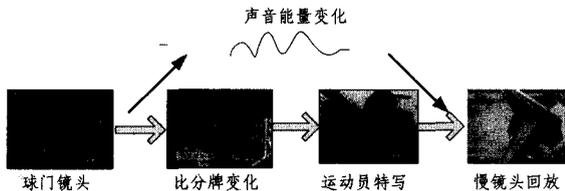


图7 点球语义概念实例

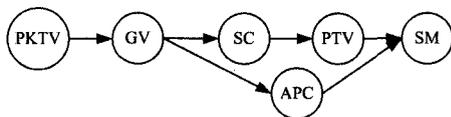


图8 点球的视频语义图

此外,可对其他类型体育视频定义图形建模元素表示的语义内容。比如网球比赛视频,网球得分事件可以描述为“球员发球特写”——“远景镜头”——“欢呼声”——“慢镜头”——“比分牌变化”,这个语义事件可以用VSG描述,如图9所示。

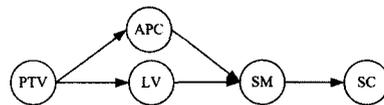


图9 网球得分的视频语义图

在视频语义图中,两种图形元素分别表示语义内容和语义之间的关系。语义内容图形元素之间的关系是视频中的时间关系和空间关系,而语义关系图形符号则表示这种时间关系和空间关系。由上例,可以归纳出图模型的图形元素的位置关系与图模型具体描述表示的视频语义内容有紧密的关系。在某种意义上可以认为,这种位置关系是由视频语义内容决定的。

#### 4.2 足球视频语义图的数据结构

分析语义图的结构不难发现,一般情况下,视频语义图(Video Semantic Graph)是简单图,并且图中结点数和边数较少,考虑到图的邻接表结构,其适合表示视频语义图。然而,要判定任意两个结点之间是否有边相连,则需搜索链表。因此,计算结点的入度和出度,不及邻接矩阵方便。十字链表是有向图的另一种链式存储结构。若将有向图的邻接矩阵看成是稀疏矩阵,则十字链表也可以看成是邻接矩阵的链表存储结构。本文提出使用十字链表(Orthogonal List)存储结构,其中定义了一个枚举类型 NodeType,表示结点的类型信息,描述结点代表的视频语义内容。

#### 5 基于DFS的视频语义图匹配算法

提出的视频语义图(VSG)主要有两个特点:一是能清晰地表示视频语义内容,并且结构简单;二是模型的图形元素,特别是语义内容元素的图形符号形状多样,有利于计算机区分不同的语义内容事件。由于VSG的这两个特点,利用图匹配的方法实现视频的检索是可行的。

把算法的适用范围限制在具有清晰编辑特点的足球比赛视频检索方面,结合提出的视频语义图(VSG)的建模元素的特殊性,提出了基于DFS的视频语义图匹配算法。

一般来说,处理的数据量越大,处理的复杂度就越大。视频包含十分丰富的数据,处理起来异常复杂。因此,在处理视频数据之前,一般的做法是将视频分割成一个又一个的镜头。用专业的软件进行视频分割,例如IBM公司的VideoAnnEx,可以方便地将视频按镜头分割。本文也采取这种思路,先将视频按镜头分割,以减少视频语义图的结点数,减小算法的复杂度。

##### 5.1 算法思路

首先,对视频场景构造视频语义图(VSG1),对事件也构造视频语义图(VSG2)。之后,遍历VSG1和VSG2。本文的算法遍历VSG1的结点,利用十字链表可以记录结点 $V_i$ 的入度( $deg^+(V_i)$ )和出度( $deg^-(V_i)$ )。再设计一个字符数组,记录结点 $V_i$ 父结点的类型和个数(入度、出度),与VSG2进行比较。若VSG1和VSG2中,相同结点的入度和出度相等,并且父结点的类型也相同,则在VSG1中存在与VSG2同构的子图。即在VSG1表示的视频中,检索到了输入图表示的视频语义内容相同的视频,完成了检索。否则,不存在。

##### 5.2 算法步骤

算法步骤如下:

步骤 1 假定针对视频构造的视频语义图为 VSG1 和检索事件的视频语义图为 VSG2。

步骤 2 调用深度优先搜索算法 (DFS) 遍历 VSG1 和 VSG2。分别计算 VSG1 和 VSG2 的结点的入度和出度,并用结点信息表 VexInfo1 和 VexInfo2 记录各个结点的出(入)度、类型和结点的父结点的类型。

步骤 3 依据 VSG2 的结点信息表 VexInfo2,针对每个结点查找 VSG1 的结点信息表 VexInfo1 中是否有相同类型的结点。若有,则转入步骤 4。否则,算法停止,返回结果,未检索到。

步骤 4 在 VexInfo2 中,针对每个结点比较 VexInfo1 中相应结点的类型、出(入)度和父结点的类型。若都相同,则转入步骤 5。否则,算法停止,返回结果,未检索到。

步骤 5 返回结果,检索到视频,播放相应的视频。

其中,结点信息表 VexInfo 的结构如下:

```
Struct VexInfo
{ NodeType;
  Deg+(V);
  Deg-(V);
  FatherNodeType;
  ChildNodeType; }
```

需要注意的是,NodeType 是一个枚举变量,用整数表示结点类型。在结点信息表 VexInfo 中,若结点  $V_i$  没有父结点或子结点,即入度或者出度为 0,则相应的栏填入  $\infty$ ,且  $\infty$  与所有的值匹配。若结点  $V_i$  的入(出)度大于 1,则在 VexInfo 中记录入(出)度数量次,并分别写入对应的父结点的类型。一般情况下,入(出)度不会超过 2,且通常是入度大于 1。

基于本文提出的算法,必须对 DFS 函数<sup>[6]</sup>做改进,以得到结点信息表 VexInfo。改进后的 DFS 函数如下:

```
//——改进的 DFS 函数——
#define int VexInfo[MAX][5]; //二维数组 VexInfo[MAX * 5] 表示结点信息表 VexInfo
void DFSTraverse(Graph G, Status (* VisitFunc)(int v)) {
    //对图 G 作深度优先遍历
    VisitFunc = Visit; //使用全局变量 VisitFunc
    for(v=0; v<G.nodenum; ++v) visited[v]=FALSE; //访问标志数组初始化
    for(v=0; v<G.nodenum; ++v)
        if (! visited[v]) {
            //对尚未访问的结点调用 DFS,并记录结点信息,写入 VexInfo
            [MAX][5]
            DFS(G, v);
            VexInfo[v][0]= NodeType;
            VexInfo[v][1]= Deg+(V);
            VexInfo[v][2]= Deg-(V);
            VexInfo[v][3]= FatherNodeType;
            VexInfo[v][4]= ChildNodeType;
        }
}
```

### 5.3 算法复杂度分析

提出的基于 DFS 的视频语义图匹配算法,对结点类型给出了恰当的约束。并且由于足球视频本身的特点,所构造的视频语义图的结点数少,结构都比较简单,因此本算法的时间复杂性主要取决于 DFS 的时间复杂性和结点比较的时间复杂性。

深度优先搜索在遍历图时,对图中每个结点至多调用一

次 DFS 函数。因为一旦某个结点被标志成已被访问,就不再从它出发进行搜索。因此,遍历图的过程实质上是对每个结点查找其邻接点的过程。其耗费的时间则取决于所采用的存储结构。当以邻接表作为图的存储结构时,找邻接点所需的时间为  $O(e)$ ,其中  $e$  为图中有向边的个数。由此,当以邻接表作为存储结构时,深度优先搜索遍历图的时间复杂度为  $O(n+e)$ 。本文采用十字链表作为图的存储结构,深度优先搜索遍历图的时间复杂度与以邻接表作存储结构时相同,也是  $O(n+e)$ 。

由于采用了二维数组表示结点信息表 VexInfo,相互比较两个二维数组的对应元素的时间复杂度最大是  $O(n^2)$ ,因此本算法的时间复杂度是  $O(n^2)$ 。

本文提出的模型使用十字链表作为存储结构,因此空间复杂度最大是  $O(n^2)$ 。

## 6 算法应用与结果分析

通过一个示例进一步解释本文提出的算法应用过程。例如,一段足球视频中,开球后一方向另一方进攻,前锋带球至禁区前,射门,进球。这段视频的视频语义图 (VSG) 如图 10 所示。

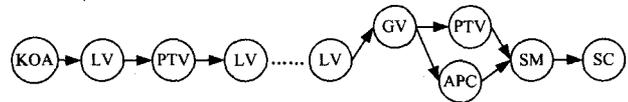


图 10 本视频语义图

在处理之前,利用 VideoAnnEx 把这段视频分割,得到一个包含进球的视频段,其视频语义图 (VSG1) 如图 11 所示。进而得到基于图模型定义的“进球”的视频语义图 (VSG2),如图 12 所示。

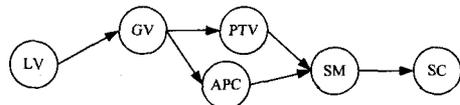


图 11 包含进球的视频语义图

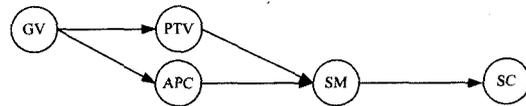


图 12 进球的视频语义图

由本文提出的算法,遍历 VSG1 和 VSG2 后得到的 VexInfo1 和 VexInfo2,如图 13 所示。

7	$\infty$	1	$\infty$	6	6	$\infty$	2	$\infty$	0
6	1	2	7	0	6	$\infty$	2	$\infty$	5
6	1	2	7	5	0	1	1	6	1
0	1	1	6	1	5	1	1	6	1
5	1	1	6	1	1	2	1	0	8
1	2	1	0	8	1	2	1	5	8
1	2	1	5	8	8	1	$\infty$	1	$\infty$
8	1	$\infty$	1	$\infty$					
VexInfo1					VexInfo2				

图 13 VexInfo1 和 VexInfo2

依据 VexInfo2 按行进行比较,在 VexInfo1 查找到了相同的部分,则说明检索成功。

(下转第 297 页)

问题,全局运动区域中还存在一些前景运动物体像素。我们下一步的研究工作是通过全局运动区域的一些辅助处理,如运动矢量场的后处理、时空联合分析等措施来进一步净化全局运动区域,从而得到更为精确的全局运动估计结果。

### 参考文献

[1] 郭丽,龚声蓉,崔志明.一种快速全局运动估计算法[J].中国图形图像学报,2007,12(12):2119-2126

[2] 李乔亮,汪国有,张桂林,等.基于掩模金字塔的高精度全局运动估计算法[J].计算机辅助设计与图形学学报,2009,21(6):758-762

[3] Arvanitidou M G, Glantz A, Krutz A, et al. Global motion estimation using variable block sizes and its application to object segmentation[C]//Proceedings of Image Analysis for Multimedia Interactive Services. 2009:173-176

[4] 彭小宁,杨明,邹北骥,等.基于局部图金字塔的不规则块匹配视频分割方法[J].计算机科学,2008,4:233-237

[5] Zheng Jia-li, Ji Xiang-yang, Ni Guang-nan, et al. Extended Direct Mode for Hierarchical B Picture Coding [J]. Proceedings of IEEE International Conference on Image Processing (ICIP), 2005,2(1):265-268

[6] Press W H, et al. Numerical Recipes in C: the Art of Scientific Computing[M]. Cambridge: Cambridge University Press, 2005

[7] Zheng Jia-li, Shen Yan-fei, Zhang Yong-dong, et al. Adaptive Selection Motion Model for Panoramic Video Coding[C]// Proceedings of IEEE International Conference on Multimedia & Expo (ICME), 2007:1319-1322

[8] Rath G B, Makur A. Iterative least squares and compression based estimations for a four-parameter linear global motion model and global motion compensation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 9(7): 1075-1099

[9] Li W P, et al. MPEG-4 Video Verification Model version 16. 0, ISO/IEC/JTC1/SC29/WG11, Doc. No. N3312[S]. Netherlands, 2006

[10] Albu F, Florea C, Zamfir A, et al. Low Complexity Global Motion Estimation Techniques for Image Stabilization[C]// Proceedings of International Conference on Consumer Electronics, 2008:1-2

[11] Bin Q, Ghazal M, Amer A. Robust Global Motion Estimation Oriented to Video Object Segmentation[J]. IEEE Transactions on Image Processing, 2008,17(6):958-967

[12] Gonzales R C, Woods R E. Digital Image Processing[M]. Addison-Wisley, 2005

[13] Chen Y M, Bajic I V. Motion Vector Outlier Rejection Cascade for Global Motion Estimation[J]. IEEE Transactions on Signal Processing, 2009,17(2):197-200

(上接第 269 页)

本文定义了“进球”、“点球”和“角球”3类视频语义图。实验数据如表2所列。

表2 实验数据集

序号	比赛名称	长度	机构	时间
1	德国 VS 波兰	1 小时 35 分	CCTV5	2006
2	瑞典 VS 巴拉圭	1 小时 37 分	CCTV5	2006
3	意大利 VS 美国	1 小时 33 分	东方卫视	2006
4	巴西 VS 日本	1 小时 34 分	东方卫视	2006

实验结果如表3所列,从表中可以看出查找到率都高于查准率,而且平均准确率大于85%。这说明采用VSG描述的五类足球视频语义是准确的,且本文提出的算法实用可行。

表3 实验结果

足球语义	实际包含数量	正确匹配数量	错误匹配数量	查准率 (%)	查到率 (%)
进球	10	10	3	76.9	100
角球	40	34	0	85	85
点球	2	2	0	100	100
平均值				87.3	95

分析实验结果发现,造成“进球”匹配错误主要是由视频中出现的精彩射门引起的。而“角球”的正确匹配数量小于实际数量的原因是比赛视频中,有些角球镜头没有给出角球区镜头。相比之下,“点球”的匹配准确率最高,这是因为点球的情况较少,并且描述点球的视频一般比较规范,受其他情况影响较少。

采用简单线性加权的融合方法<sup>[4]</sup>对测试数据进行实验,查准率和查全率的平均值分别是42%和83%,可见本文的方法明显优于基于简单线性加权的融合方法。

**结束语** 提出了新的图模型的视频语义内容建模方法,构建了视频语义图(VSG)。从3个方面详细地阐述了图模型,第一,如何定义每个建模元素的图形符号;第二,如何定义

图形符号之间的位置关系;第三,如何将表示法映射到抽象语义。针对足球视频,对模型进行了深入的讨论,提出了基于图模型的足球视频语义内容建模方法。可以将其推广到其他视频语义内容建模,允许用户自行定义图形元素表示的语义。在分析了图匹配的相关问题之后,提出了基于DFS的图模型的视频检索算法,详细阐述了算法思路和步骤,分析了算法的复杂度,并用实验验证了算法的可行性。

### 参考文献

[1] Jain R. Multimedia Information Retrieval; Watershed Events. MIR'08[C]// Vancouver, British Columbia, Canada, October 2008:30-31

[2] Rui Y, Gupta A, Acero A. Automatically extracting highlights for TV baseball programs[C]// ACM Multimedia, 2000

[3] Wang Feng, Jiang Yu-gang, Ngo C-W. Video Event Detection Using Motion Relativity and Visual Relatedness[C]// ACM MM'08. Vancouver, British Columbia, Canada, 2008

[4] 白亮. 本体支持的视频情报分析方法与技术研究[D]. 长沙:国防科技大学, 2008

[5] 李刚. 知识发现的图模型方法[D]. 北京:中国科学院软件研究所, 2001

[6] 严蔚敏,吴伟明. 数据结构[M]. C语言版. 北京:清华大学出版社, 1997

[7] Goldberg A B, Li M, Zhu X. Online manifold regularization: A new learning setting and empirical study[C]// ECML PKDD, 2008

[8] Aly R, Hiemstra D. Concept Detectors: How Good is Good Enough[C]// MM'09. Beijing, 2009

[9] Han Bo, Yan Yan, Chen Zhen-ghua, et al. A General Framework for Automatic On-line Replay Detection in Sports Video[C]// MM'09. Beijing, 2009

[10] 侯绿林,白亮,老松杨.一种压缩域中的体育视频慢镜头探测方法[J].计算机科学,2009,36(9):283-286