

# 基于运动特征的视频检索

程照辉 毋立芳 刘健

(北京工业大学电子信息与控制工程学院, 北京 100124)

**摘要:** 运动特征是描述视频内容的一种重要的信息,也是视频区别于图像数据所特有的内容,要想对视频内容进行全面的刻画,运动特征是必不可少的一个方面。本文提出了一种基于时空运动特征的视频检索方案。根据离散的球的位置,获取台球原始运动轨迹,进而修正得到完整的台球运动轨迹;以运动轨迹点序列作为运动特征,同时提出了基于运动轨迹点序列顺序匹配的视频段分步检索方法:首先根据碰撞点将运动轨迹分段,然后对检索视频的轨迹分段重采样,最后利用轨迹模型匹配方法实现检索。实验结果表明,我们提出的方案实现了一定程度的语义检索,能够较好地检索到符合人们主观理解的斯诺克运动轨迹的视频内容。

**关键词:** 运动特征; 视频检索; 运动轨迹; 重采样

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 1003-0530(2011)05-0765-06

## Motion feature based video retrieval

CHENG Zhao-hui WU Li-fang LIU Jian

(School of Electronic Information and Control Engineering, Beijing University of Technology, Beijing, 100124)

**Abstract:** Motion feature is one of important clues representing content of a video sequence, and also it is the typical content that distinguished video data from image data. It is an essential aspect and while extracting the comprehensive description of the video content. In this paper we propose a spatial-temporal motion feature based video retrieval scheme. The initial trajectory of balls is obtained, according to the discrete location of the ball. And then the trajectories are modified by the collision point. The trajectory dots sequence is considered as feature of motion, at the same time, we propose a video clip retrieval method by step which is based on the trajectory dots sequence order match: firstly, the trajectory is divided up into several sections, according to the collision point, and then the trajectory segments of the video are re-sampled, finally, the trajectory model matching method is used to achieve retrieval. The experimental results show that the proposed scheme could query exactly and achieve semantic content retrieval to some extent, it provides the retrieval result consistent with the snooker video's moving trajectory content of people's subjective understanding.

**Key words:** motion feature; video retrieval; trajectory; resample

## 1 引言

随着多媒体技术的迅速发展,基于内容的视频检索得到了越来越广泛的应用。一般认为,视频数据所包含的视觉内容包括颜色、纹理、形状和运动的信息<sup>[1]</sup>,其中运动信息是视频区别于图像数据所特有的内容<sup>[2]</sup>。对于一个视频片段来说,运动信息是反映视频中变化演进的重要信息<sup>[3]</sup>。很多视频检索系统采用颜色、纹理、形状等图像的低级特征来进行查询,但是这些低级特征不能充分代表视频中丰富的语义信息,而且人们习惯根据对象行为(action)描述视频中的语义内容。因此基于运动信息的检索可以提供与颜色、纹理无关的检索特征,要想对视频内容进行全面的刻

画,运动信息是必不可少的一个方面。这在场景相对单一的体育运动类视频检索、数字监控等系统中有应用的需求。基于运动的视频检索揭示了视频内容的动态特征,成为视频检索的活跃领域。本文也正是在这样的背景下对此进行了初步的研究和实验。

运动轨迹由视频序列中运动对象的质心位置组成,是视频运动特征中重要的一种特征,它反映了视频对象运动的时间-空间关系<sup>[4]</sup>。Man-Kwan Shan<sup>[5]</sup>等,提出了基于运动轨迹的检索方案,通过把轨迹看作是一些线段的序列的方法来对轨迹建模,其中每一个线段用它们的倾斜度来表示。简化了轨迹的表示,节约了响应时间。他们还提出了基于动态规划的相似度测量方法来进行匹配检测。VideoQ<sup>[6]</sup>将运动作为检索特征的一部分,他

采用形态分割提取运动对象,每一帧中对象的坐标被提取出来,得到对象的轨迹矢量后结合其他特征进行匹配检索。A. Dyana<sup>[7]</sup>等,提出了基于形状特征和运动轨迹相结合的检索方案,其中形状特征用曲率尺度空间表示,运动轨迹用多项式曲线拟合表示。Sahouria<sup>[8]</sup>分析了从监视应用场景中得到的视频。他在获得了对象运动轨迹的基础上,将轨迹用两个运动矢量表示,即:  $X$  和  $Y$ ,而后分别对其进行小波变换,然后提取出前八个变换系数作为运动特征向量。S. Dagtas<sup>[9]</sup>通过对运动建模,利用时空相关性的特征,建立了几种具有绝对平移等特征的检索方案,最终建立了一个实验模型。文献[10]提出了基于运动区域的轨迹分析与重排的视频检索方案。文献[11]建立了视频自动分析系统,其中包括基于轨迹的视频注释功能。文献[2][12]针对特征点提出了基于特征点序列的顺序匹配检索。

针对斯诺克台球的运动轨迹的特点,本文提出了一种基于运动轨迹点序列顺序匹配的视频分步检索方法。在场景相对单一的斯诺克运动视频,利用一种有效的多目标检测方法提取视频中对象的坐标,获取斯诺克运动视频中台球对象的运动轨迹。由于视频采集的硬件环境要求不高,采集得到的球和球的碰撞点、球和边库的碰撞点经常采集不到,因此需要对运动轨迹进行修正,获取完整的台球运动轨迹。根据碰撞点将运动轨迹分段,并结合检索视频进行分段重采样,得到顺序序列轨迹采样点,最后根据序列轨迹采样点的匹配结果进行检索。

本文后续内容安排如下:第二部分介绍视频检索的方案框架,第三部分介绍斯诺克运动中台球的运动轨迹修正,第四部分介绍了基于轨迹点序列顺序匹配的分步检索,第五部分给出实验结果,最后总结本文。

## 2 方案框架

基于运动特征的视频检索的方案框架,如图1所示。

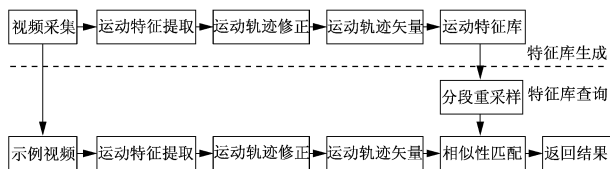


图1 基于运动特征的视频检索方案框架

Fig. 1 Framework of video retrieval based on motion feature

在方案中,使用视频采集卡采集场景相对单一的斯诺克运动视频,利用一种有效的多目标检测方法提取视频的运动特征,根据运动特征中的对象坐标提取对象的运动轨迹及其矢量,并将其作为视频序列的运动特征存入特征库;用户检索可以采用示例查询机制,根据示例运动轨迹及其矢量特征,并结合检索视频进行分段重采样,

得到新的轨迹采样点。通过特征匹配,在特征库中查找最相似的视频片段,并依据相似度进行排序,最后按照排序的结果从视频库中取出视频,返回给用户。

方案中运动特征提取即台球的检索和运动跟踪,确定各个台球在台面上的准确坐标位置,连续帧的台球坐标位置变化,就体现了视频中台球的运动特征。文献[13]针对斯诺克运动视频,采用一种有效的多目标检测方法,检测斯诺克视频中的台球,确定了不同颜色球在台面上的准确坐标。

## 3 运动轨迹修正

系统采用普通摄像头采集视频(每秒25帧),这样实际得到的是离散图像序列,球和球的碰撞点、球和边库的碰撞点经常采集不到,图5是一个母球碰撞红球,红球落袋的视频,从轨迹上可以看出,母球与红球的碰撞点、母球和边库的碰撞点均没有采集到。因此需要对台球的原始运动轨迹进行修正:首先根据相邻帧离散的台球位置,描绘出台球运动的原始运动轨迹;然后根据球的运动轨迹计算碰撞点的位置,对台球运动轨迹进行校正、调整、描绘出台球运动的完整运动轨迹。

首先获取一杆斯诺克台球运动的连续帧图像,确定各帧图像中各台球的台面平面坐标位置,具体算法步骤如图2所示:

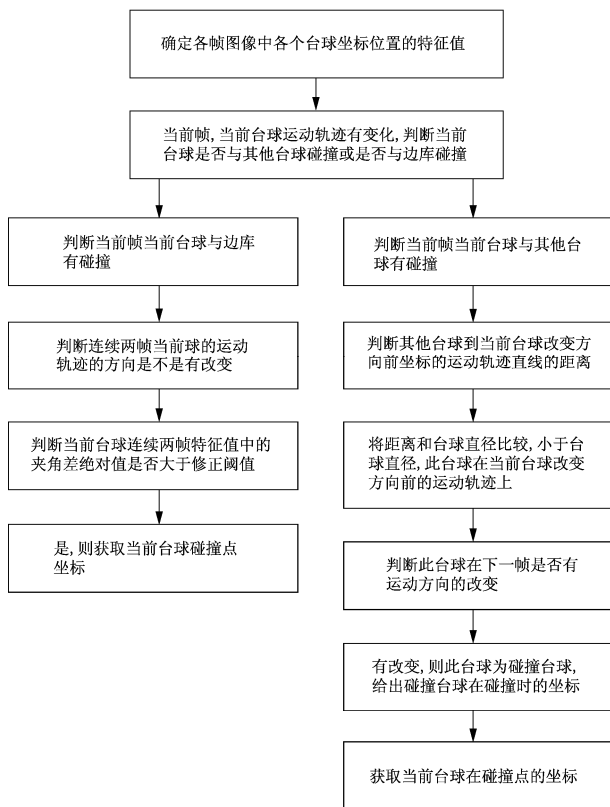


图2 台球的运动轨迹修正示意图

Fig. 2 Flow chart of billiard trajectory correction

其中:计算各帧图像中各个台球的坐标位置点的特征值,所述特征值包括:各个台球当前帧坐标点与前一帧坐标点构成的直线针对 X 轴的斜率、两坐标点的 X 坐标之间的差值、Y 坐标之间的差值、两坐标点的直线距离、直线与 X 轴正方向的夹角;

球与球碰撞时修正碰撞点的坐标如图3所示:设:当前台球2为白球,台球3为碰撞球,虚线台球4为虚拟的碰撞时的白球2,当前台球的运动轨迹的直线 A 的直线方程为 line1,由碰撞前的运动轨迹的斜率计算出运动轨迹直线的垂线的斜率,经过碰撞球的垂线 B 的直线方程 line2,计算两条直线的交点 o (垂点),从而计算当前球当前点(a)到垂点(o)的距离 |ao| 以及碰撞球的坐标(b)到垂点(o)的距离 |bo|,则碰撞点(c)到垂点(o)的距离为 |co|,进而计算 c 点的坐标即为碰撞点的坐标;

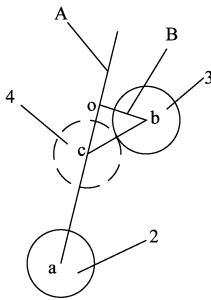


图3 修正球与球碰撞点的坐标示意图

Fig. 3 Modify the coordinates of the collision point when two balls collided

修正球与边库碰撞点的坐标如图4所示:由特征值中当前点 j 的斜率和下一点 k 的斜率,分别计算经过当前点的直线方程 f1 和经过下一点的直线方程 f2,计算两条直线的交点 f,图中的虚线 m 是未修正前的轨迹;如果交点的 x 坐标为正数且小于长边的长度, y 坐标为正数且小于短边的长度,则此交点即为碰撞点。

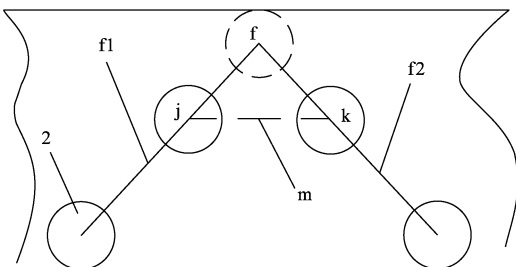


图4 修正球与边库碰撞点的坐标示意图

Fig. 4 Modify the coordinates of the collision point when a ball collided with the board's side

经过上面的分析后可以获取斯诺克运动视频中,台球的未修正的运动轨迹和修正后的运动轨迹如图5和6所示:

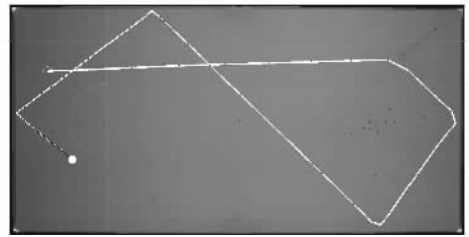


图5 未修正时,白球碰撞红球,红球落袋的轨迹图

Fig. 5 The initial trajectory of balls

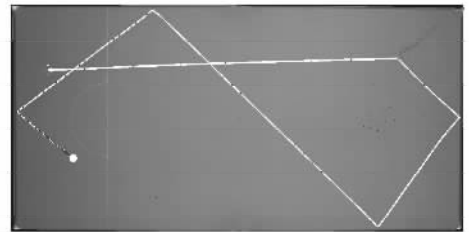


图6 修正后,白球碰撞红球,红球落袋的轨迹图

Fig. 6 The modified trajectory of balls

## 4 基于运动轨迹点序列顺序匹配分步检索方法

为了处理用户给出的示例视频段,必须在运动的表示之间建立一个较好的相似性匹配测量机制,这种测量机制必须基于运动的某种或某些特征<sup>[12]</sup>,从而可以满足用户的检索需求。针对文献[2][12]中提出的基于特征点顺序匹配检索,本文根据台球的运动轨迹的特点对此方法进行了改进,实现了基于运动轨迹点序列顺序匹配分步检索的方法,首先根据碰撞点将运动轨迹分段,利用轨迹段数进行粗略检索,快速检索出和示例视频具有相同的轨迹段数的视频片段;其次,把检索视频分段重采样,计算新的台球的轨迹采样点。最后利用轨迹模型匹配方法实现准确检索。

### 4.1 分段重采样

用户给出样例视频,寻找数据库中的相似性匹配的视频数据时,根据碰撞点将运动轨迹分段,利用轨迹段数进行粗略检索,快速检索出和示例视频具有相同的轨迹段数的视频片段。然后分别计算示例视频轨迹中每一段的长度,起始坐标点和同一轨迹段中第二个坐标点的长度,计算两个长度距离的数值比。对于检索的视频段轨迹点来说,计算起始点距离该轨迹段第二点的长度与该轨迹段的全长的数值比。两个比值相等,则可以计算出检索的视频轨迹中的第二点的坐标,依次可重新计算检索视频段中的所有轨迹点。实现轨迹点重新采样。

假设示例视频轨迹中的一段起始点为  $(x_1, y_1)$ , 第二个点为  $(x_2, y_2)$ , 第三点为  $(x_3, y_3)$ ... 第 k 点为  $(x_k, y_k)$ , 终点为  $(x_n, y_n)$ ; 检索的视频轨迹段起始点为  $(x'_1, y'_1)$ , 第 k 点为  $(x'_k, y'_k)$ , 终点为  $(x'_n, y'_n)$ 。

则:示例视频中起始点 $(x_1, y_1)$ 到终点 $(x_n, y_n)$ 的距离为:

$$d = \sqrt{(x_n - x_1)^2 + (y_n - y_1)^2} \quad (1)$$

检索视频中起始点 $(x'_1, y'_1)$ 到终点 $(x'_n, y'_n)$ 的距离为:

$$d' = \sqrt{(x'_n - x'_1)^2 + (y'_n - y'_1)^2} \quad (2)$$

示例视频中起始点 $(x_1, y_1)$ 到第 $k$ 点 $(x_k, y_k)$ 的距离为:

$$d_k = \sqrt{(x_k - x_1)^2 + (y_k - y_1)^2} \quad (3)$$

检索视频中起始点 $(x'_1, y'_1)$ 到第 $k$ 点 $(x'_k, y'_k)$ 的距离为:

$$d'_k = \sqrt{(x'_k - x'_1)^2 + (y'_k - y'_1)^2} \quad (4)$$

由 $d_k/d = d'_k/d'$ 推出:

$$d_k/d = \sqrt{(x'_k - x'_1)^2 + (y'_k - y'_1)^2} / d' \quad (5)$$

检索视频中的轨迹直线的斜率为:

$$k' = (y'_n - y'_1) / (x'_n - x'_1) \quad (6)$$

把公式(6)代入(5)中,推出:

$$d_k/d = \sqrt{(1+k'^2)} |x'_k - x'_1| / d'$$

则可计算出第 $k$ 点的 $x$ 坐标:

$$\begin{cases} x'_k = x'_1 + d_k d' / (d \sqrt{1+k'^2}), x'_k \geq x'_1 \\ x'_k = x'_1 - d_k d' / (d \sqrt{1+k'^2}), x'_k < x'_1 \end{cases} \quad (7)$$

对于检索视频中起始点与第 $k$ 点的斜率为:

$$k' = (y'_k - y'_1) / (x'_k - x'_1) \quad (8)$$

把式(8)代入(7)中,得第 $k$ 点的 $y$ 坐标:

$$y'_k = y'_1 + k' (x'_k - x'_1)$$

$$= \begin{cases} y'_1 + k' d_k d' / (d \sqrt{1+k'^2}), x'_k \geq x'_1 \\ y'_1 - k' d_k d' / (d \sqrt{1+k'^2}), x'_k < x'_1 \end{cases}$$

其中 $k=2, 3, \dots, n-1$ ,则对于检索的视频片段可以分段重采样,计算出新的轨迹点的坐标。

#### 4.2 轨迹模型匹配方法

一段轨迹的所有点对应于一个 $n$ 维向量,它所对应的不过是 $n$ 维空间里的一个点。同样道理,对于一个特征向量,从空间的角度来看,它也不过是特征空间里的一点而已,因此我们得到的特征向量都可以看作我们的特征空间里的点,所谓的查询过程演化成了寻求满足条件的点的过程。

设 $C(k)=[x(k), y(k)], k=1, 2, \dots, N$

数组 $C$ 有两个元素,分别代表视频中对象的纵横坐标, $N$ 为视频序列的总数目。显然, $N$ 个点定义了屏幕上运动对象的轨迹,并且用来进行查询视频段( $C^Q$ )

和视频数据库中视频数据( $C^D$ )之间的相似度测量。对于这种轨迹模型的匹配处理可表示如下方式。

对于存储在数据库中的各元素( $C^D$ ),如果 $Dist(C^D, C^Q) < T$ ,那么就接受 $C^D$ ,其中, $C^Q(k)=[q_x(k), q_y(k)], C^D(i)=[x(i), y(i)], k=1, 2, \dots, N_q, i=1, 2, \dots, N, N_q$ 是查询序列的帧数目, $q_x, q_y$ 指定了 $x$ 轴和 $y$ 轴的中心点的轨迹。

作为一个通用的相似度测量方案,我们使用欧氏距离来进行视频序列的匹配。这样,有不相似度测量函数( $Dis()$ )和距离测量函数( $Dist()$ )。

$$Dis(C^D, C^Q) = \min_i [Dist(x_q^i, q_x) + Dist(y_q^i, q_y)] \quad (9)$$

$$Dist(x_q^i, q_x) = (x_q^i - q_x)^2 (x_q^i - q_x) = \mathbf{x}_q^{i,T} \mathbf{x}_q^i + 2\mathbf{x}_q^{i,T} q_x + q_x^T q_x \quad (10)$$

$$Dist(y_q^i, q_y) = (y_q^i - q_y)^2 (y_q^i - q_y) = \mathbf{y}_q^{i,T} \mathbf{y}_q^i - 2\mathbf{y}_q^{i,T} q_y + q_y^T q_y \quad (11)$$

其中 $q_x$ 和 $q_y$ 是查询序列, $x_q^i$ 和 $y_q^i$ 是具有与 $q_x$ 和 $q_y$ 等长的、 $x$ 和 $y$ 的一部分,也就是 $x_q^i(k) = x(k+i), k=1, 2, \dots, N_q; i=0, 1 \dots N-N_q-1; x, y \in Z^N; q_x, q_y, q_z, x_q, y_q \in Z^{N_q}; N_q \leq N$ 。 $Dist(x_q^i, q_x)$ 、 $Dist(y_q^i, q_y)$ 各自表示查询矢量和与其等长的库中特征矢量的相应坐标上的欧氏距离。

## 5 实验结果

实验视频数据是采集的斯诺克运动视频序列8000帧图像,分60杆视频数据。随机抽取一杆作为示例视频,则每一杆运动轨迹和本身的距离如图8所示,由图7观察所知,我们提出的距离计算方法得到的同示例视频的欧氏距离都很小,平均值为0.56,最大为1,这说明同示例视频都能很好的检出。

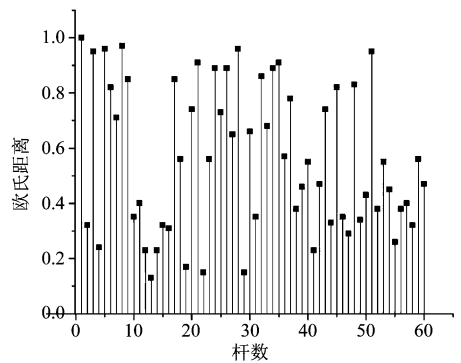


图7 60杆同示例视频的欧氏距离

Fig.7 Euclidean distance of 60 sample videos

将60杆视频按照白球的运动轨迹进行主观分类。按照欧氏距离取阈值进行同类视频检索,本文算法和文献[2][12]中算法的查全率和查准率如图8所示。

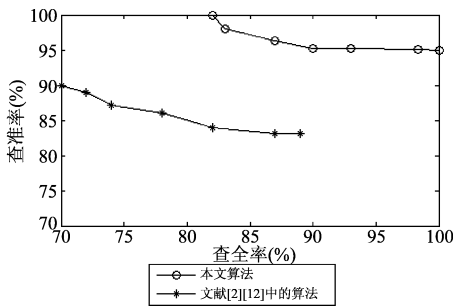


图 8 查全率和查准率  
Fig.8 Recall and precision

从图 8 可以看出,在本文算法中随着查全率的提高,查准率逐步下降。当查全率为 82% 时,算法的查准率为 100%。当查全率达到 100% 时,查准率下降到 95.2%;通过对比可知:本文提出的算法具有较好的检索性能。进一步对每一杆按照欧氏距离从小到大排序,分别取前 3 个,前 5 个,前 7 个检索结果,查看同类视频的查全率。检索结果比较如表 1 所示,从表中可以看出,本文算法的平均查全率比较高,取前 7 个检索结果,可以检索到全部同类视频。

图 9 为检索示例视频,图 10 给出检索出的前 6 个视频对应轨迹及欧氏距离。从图中可以看出,主观相似视频的距离比较小,主观不相似视频的距离比较大,这说明了本文提出的距离定义能够很好的反映视频之间的主观相似度。

表 1 检索结果

Tab.1 The result of retrieval

	文献[2][12]中的算法		本文算法	
	正确检出同类视频数	平均查全率	正确检出同类视频数	平均查全率
Top3	54	90%	57	95%
Top5	56	93.3%	59	98.3%
Top7	59	98.3%	60	100%

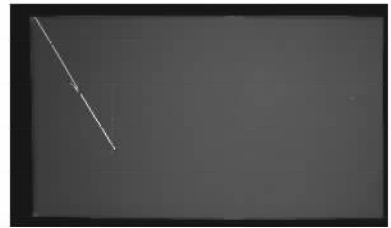


图 9 示例视频 白球碰撞粉球,粉球落袋  
Fig.9 Trajectory of sample video

通过基于运动轨迹点序列顺序匹配的分步检索方法,计算的相似度即欧氏距离最小时,即代表所对应的视频段和所提交的样例视频(也叫视频模板)最为相似,在本实验中,虽然,样例视频是取自视频源的第 12 杆数据中的 106 帧,理论上,样例视频和视频源的第 12 杆数据的欧氏距离应为 0,但实际上由于噪声的影响和算法本身存在的问题,导致了一定的误差,实验中误差值约为 0.17,但这并不影响检索的结果。

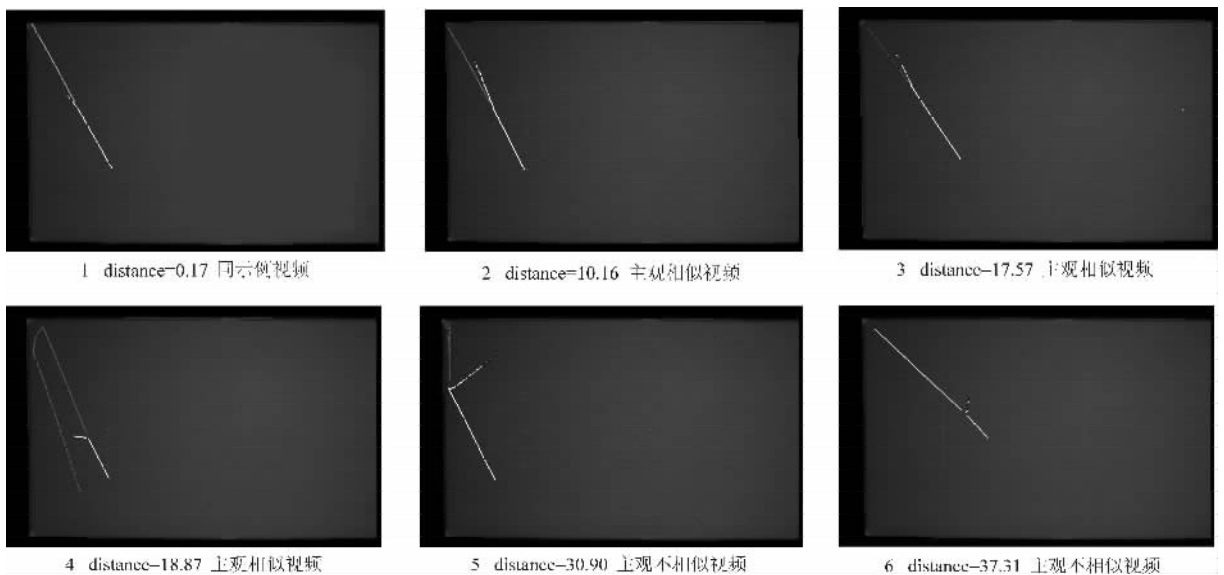


图 10 由图 9 的示例视频检索结果排序  
Fig.10 Results of searching sample video

## 6 结论

运动信息是视频数据中反映其内容变化演进的特

有信息,对于描述视频的内容具有不可替代的作用。本文根据斯诺克视频中台球的运动轨迹反映视频主要运动内容这一特点,根据视频中离散的台球的位置,获

取台球原始运动轨迹,进而修正得到完整的运动轨迹。以运动轨迹点序列作为运动特征,同时提出了基于运动轨迹点序列顺序匹配的视频段分步检索方法。最后,用8000帧共60杆数据对算法进行了测试实验。实验结果表明,基于运动轨迹点序列顺序匹配的分步检索方法,能够较好地检索到运动轨迹相符合的视频,实现了一定程度的语义检索。

#### 参考文献

- [1] Ohm J R, Bunjamin F, Liebsch W, et al. A set of visual feature descriptors and their combination in a low-level description scheme [J]. *Signal Processing: Image Communication*, 2000, 16:157-179.
- [2] 俞天力, 章毓晋. 基于全局运动信息的视频检索技术 [J]. *电子学报*, 2001, 29(12):1794-1798.  
Yu Tian-li, Zhang Yu-jin. Video Retrieval Based on the Global Motion Information [J]. *Acta Electronica Sinica*, 2001, 29(12):1794-1798. (in Chinese)
- [3] Minh-Son Dao, Francesco G. B. DeNatate, Andrea Massa. Video Retrieval using Video Object-Trajectory and Edge Potential Function [C]. *International Symposium on Intelligent Multimedia, Video and Speech Processing, ISIMP 2004*: 454-457.
- [4] Bashar Tahayna, Mohammed Belkhatir, Saadat Alhashmi. Motion Information for Video Retrieval [C]. *IEEE International Conference on Multimedia and Expo, ICME 2009*, p 870-873.
- [5] Man-Kwan Shan, Suh-Yin Lee. Content-based Video Retrieval via Motion Trajectories [C]. *In Processing of the International Conference on SPIE*, 2000, Vol. 3561, pp. 52-61.
- [6] S.-F. Chang, W. Chen, H. J. Meng, H. Sundaram, D. Zhong. A Fully Automatic Content Based Video Search Engine Supporting Multi-Object Spatio-temporal Queries [J]. *IEEE Transactions on Circuits and Systems for Video Technology. Special Issue on Image and Video Processing for Interactive Multimedia*, 1998, 8(5):602-615.
- [7] A. Dyana, M. P. Subramanian, Sukhendu Das. Combining Features for Shape and Motion Trajectory of Video Objects for Efficient Content based Video Retrieval [C]. *Seventh International Conference on Advances in Pattern Recognition*, 2009, p113-116.
- [8] E. Sahouria. Video indexing based on object motion. M. S. thesis, Dept. of EECS, UC Berkeley, Berkeley, CA. 2000.

- [9] S. Dagtas, W. Al-Khatib, A. Ghafoor, R. L. Kashyap, Models for Motion Based Video Indexing and Retrieval [J]. *IEEE Transactions on Image Processing, Special Issue on Image and Video Processing for Digital Libraries*, January 2000, 9(1):88-101.
- [10] Feng Bailan, Cao Juan, Lin Shouxun, Zhang Yongdong, Tao Kun, Motion Region-based Trajectory Analysis and Re-ranking for Video Retrieval [C]. *IEEE International Conference on Multimedia and Expo, ICME 2009*: 378-381.
- [11] Li Yang, Zhang Yafei, Lu Jianjiang, Li Ran, Wang Jiabao. Video Analysis and Trajectory based Video Annotation System [C]. *Asia-Pacific Conference on Wearable Computing Systems (APWC 2010)*, 2010: 307-310.
- [12] 姚领田, 刘雨, 程勇. 一个基于运动的视频检索系统 [J]. *计算机与现代化*, 2004, 2(10):20-22.  
Yao Ling-tian, Liu Yu, Cheng Yong. A Motion-based Video Retrieval System [J]. *JiSuanJi Yu XianDaiHua*, 2004, 2(10):20-22. (in Chinese)
- [13] Wu Li-fang, Liu Jian, Cheng Zhaohui, Wang Hong, Liu Qi. An Effective Multi-object Detection Approach [C]. *International Symposium on Intelligent Signal Processing and Communication Systems*, 2010, Page(s):1-4.

#### 作者简介



程照辉(1982-),男,河南濮阳人,现为北京工业大学信息与通信工程专业硕士研究生。主要研究方向:视频分析及应用。本课题是与北京瑞盖台球科技公司合作的横向项目(斯诺克辅助裁判系统开发)中的一部分。

E-mail: chengzhaohui@emails.bjut.edu.cn



毋立芳(1970-),女,山西临猗县人,博士,硕士生导师,北京工业大学电子信息与控制工程学院教授。主要研究方向:视频分析及应用;人脸检测与识别;生物特征加密及应用;图像处理。

E-mail: lfwu@bjut.edu.cn



刘健(1988-),男,湖北荆州人,现为北京工业大学信息与通信工程专业硕士研究生。主要研究方向:视频分析及应用。E-mail: lockie@sina.cn