

视像中的摄像机运动检测

徐淑平

(清华大学 计算机系, 北京 100084)

摘要:摄像机的运动信息是视像内容检索中的一种重要特征。文中介绍了两种基于全局运动的摄像机运动检测方案。首先提取视像中的运动特征作为原始数据,然后分别使用了迭代最小二乘法和支持向量机分类的方法进行摄像机运动的检测,通过对试验结果的对比,前一种方案可更有效地完成摄像机运动类型的判断。

关键词:视像检索;支持向量机;最小二乘估计;摄像机运动检测

中图分类号:TP274+.5

文献标识码:A

文章编号:1673-629X(2007)07-0001-03

Camera Motion Detection in Video

XU Shu-ping

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: Motion information of camera is an important feature which describes video information. Two kinds of camera motion detection systems are proposed in this paper. First motion feature is extracted from video, then iterative least squares estimation (ILSE) and support vector machine (SVM) methods are used to detect camera motion. Experimental results indicate that the first way is more effective for camera motion detection.

Key words: video retrieval; SVM; least-square estimation; camera motion detection

0 引言

随着图像检索技术的日益成熟,视像检索已成为当今研究的一个重点。人们已提出了各种不同的方法来提取视像的特征和改进检索的方法,以此来达到提高视像的检索精度和查全率。然而视像不同于静止图像,视像的帧之间是有非常紧密的内容链接的,在视像中存在着其他媒体不具有的一种特征,那就是运动信息。在早期的研究中,人们多半都没有把其考虑在其中,因为它的处理费时和使用的复杂性,但是不可否认的是,运动信息的引入一定可以为视像的检索提供一个新的生机。在 TRECVID2005 中有一个新的子任务:视像的底层特征提取(即摄像机运动检测),要求各参评单位准确判断在给定的视像镜头中是否有摄像机的运动,以及属于什么类型的运动,文中就是以此为目标进行试验。

1 摄像机运动类型及视像运动特征的提取

人们在描述一个视像的内容时,通常把视像中的内容分为背景和前景两部分,因此在视像序列的运动分析中,也常把运动信息分为背景运动信息和前景运动信息。背景运动是由摄像机的运动造成的镜头内所有点的整体移动,它又被称为摄像机运动或全局运动;而前景运动是指被拍摄物体在场景中的运动,又称为物体运动或局部运动。这两种运动的性质是截然不同的,对于摄像机运动的整体性使得它可以用一组很少的模型参数来表示,而物体运动则较为复杂,需要有比较精细的方法才能完整地刻画出来。文中的任务就是要判断视像中的全局摄像机运动的类型。

对于摄像机的运动模型,MPEG-7^[1]标准从摄像机操作的角度进行了详细的规定,共有6种摄像机的运动类型。根据试验要求,只要求判断对应的视像镜头中是否存在三种类型的摄像机,它们分别是: PAN(平移)、TILT(俯仰摇拍)和 ZOOM(缩放)运动,如图1所示。

如何提取视像中的运动信息,在文献[2]中都有比较详细的分析,一般来说人们可以通过光流计算法、采用快匹配和直接从视像的压缩域提取的三种方式来进行视像的运动矢量(Motion Vector, MV)的提取。而对

收稿日期:2006-10-25

基金项目:国家自然科学基金重点项目(60135010);国家自然科学基金资助项目(60321002);国家重点基础研究(973)资助项目(2004CB318108)

作者简介:徐淑平(1972-),男,江西人,硕士研究生,研究方向为多媒体技术、人工智能;导师:林福宗,教授,研究方向为多媒体应用基础。

于 MPEG 编码的视像流,有效且快速获得运动信息的方法就是第三种方法,即利用 MPEG 编码中的已经存在的运动矢量,这也是基于运动特征的视像检索中经常使用的方法。但是在视像中,不同类型的图像(I 图、B 图、P 图)所包含的运动矢量的类型不同,而且各个参考帧与其被参考帧之间所间隔的帧数也不同,因此带来了运动矢量模值尺度上的不均等。所以在许多简单的应用中,可以只采用 P 图中的前向运动矢量来进行实验。

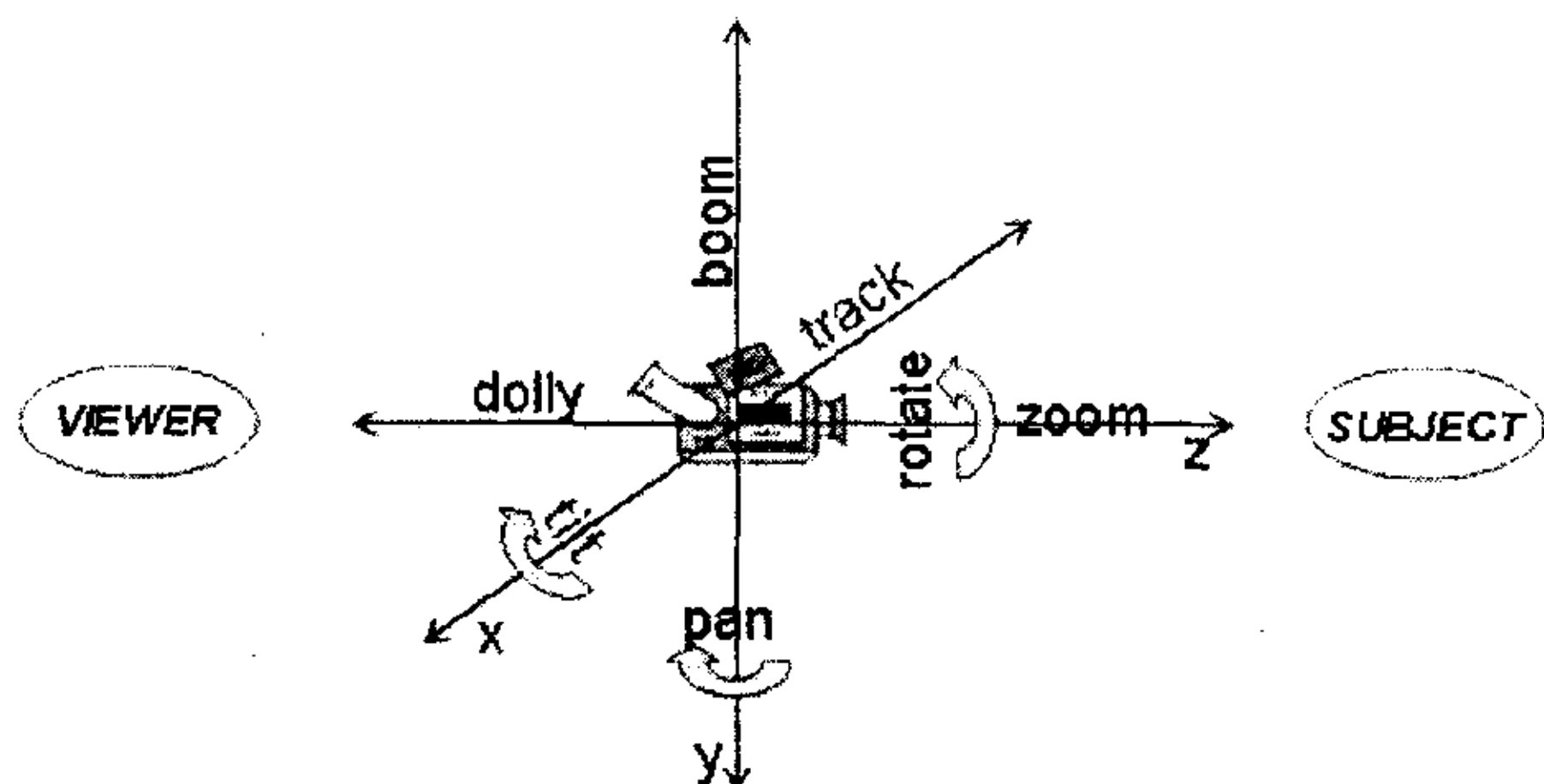


图 1 摄像机的运动类型分类

本试验采用了对每个视频帧进行 16×16 的分块划分,共产生了 20×15 个宏块,以每个宏块的中心位置为坐标点,对应生成一个运动矢量信息,并按照视像帧的顺序,把这些矢量存入到一个二进制的文件中以供调用。

2 基于仿射参数模型和迭代的最小二乘法的运动检测系统

视像中的摄像机运动体现为视像中的全局性运动,因此可以只对运动矢量进行统计分析就可以得到摄像机的运动类型,并由此来索引和检索视像。最为简单的做法是根据视像序列中帧的运动矢量方向的优势分布将摄像机运动标注为八个主要运动方向(上、下、左、右及四个斜方向)之一。另外,有的研究者也根据运动矢量是否有向焦点汇聚的特性来检测摄像机是否在摄像机的缩放运动。

对于摄像机运动进行建模既可以从相机运动的角度,也可以从参数模型的角度出发。但由于参数模型具有较好的数学形式,而且计算方便,在本实验中采用的是后者。首先建立以相机中心为原点的空间坐标系,因实验主要是提取摄像机三个方向的运动特征,故采用下面的四参数模型^[3]就可以较好的近似表达:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} p_1 & -p_2 \\ p_2 & p_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} p_3 \\ p_4 \end{pmatrix} \quad (1)$$

其中, (x, y) 和 (x', y') 分别表示相邻两帧之间对应点在各自图像中的坐标,它们是以图像中心点为坐标原点的标准正交坐标系下的数值, p_1, p_2, p_3, p_4 则表示

相机运动的四个参数。其中 p_1 表示帧间相机缩放运动参数: $p_1 > 1$ 放大, $p_1 < 1$ 缩小, $p_1 = 1$ 无缩放; p_2 代表摄像机的旋转操作,因此实验不要求此摄像机的运动,故在后面的公式中用 0 来表示 p_2 ; 比值 p_3/p_1 和 p_4/p_1 分别表示帧间相机沿垂直方向 Y 和水平方向 X 移动时的参数。

而在进行具体的摄像机运动参数的求解时,采用了迭代的最小二乘法^[4]来完成。每次使用所有不符合任何已求出的运动参数的运动矢量,进行迭代参数拟合,得到一组参数。下面是视像的每一帧的迭代最小二乘参数拟合算法的执行流程:

- ①将一帧中所有的有效宏块放入初始集合 S 中;
- ②利用参数计算公式和集合 S 中宏块的运动矢量计算相应的参数;
- ③利用公式计算出集合中每个宏块的全局运动矢量 u ;
- ④计算宏块的运动矢量 v 和计算得到的全局运动矢量 u 之间的平均误差 $\text{error}(u, v) = \sqrt{(u_x - v_x)^2 + (u_y - v_y)^2}$;
- ⑤将误差超过阈值 $\text{Threshold}_{\text{error}}$ 的宏块从集合中去掉,不再参加后面的迭代过程;
- ⑥如果平均误差小于某阈值,或者集合 S 中宏块数量少于 $\text{Threshold}_{\text{min}}$,则结束迭代。

3 基于支持向量机的运动检测

支持向量机是统计学习理论中最年轻的部分,由 Vapnik 等人于 1992 年提出^[5]。它实际上是最优超平面方法与核函数方法相结合的产物。支持向量机的工作原理是利用非线性映射将输入的特征向量映射到一个高维的特征空间中,在该空间中构造一个最优超平面来逼近目标分类函数。而最优超平面方法控制着学习机实现的函数集的容量,使得支持向量机具有良好的推广能力。另外核函数则避免了显式的非线性映射,从而克服了高维特征空间带来的计算困难。图 2 是一个基于二分类问题的 SVM 分类示意图,实心点和空心点分别表示两类样本, H 为其分类线(或面), H_1, H_2 分别为过各类中离分类线 H 最近的样本且平行于分类线的直线(或面),它们之间的距离就叫做分类间隔。所说最优分类线(或面)就是要求分类线(或面)不但能将两类无错误地分开,而且要使两类的分类间隔最大。在图 2 中 H 就达到了此目的,而 H_1, H_2 上的样本点就称作支持向量。

SVM 使用支持向量机 Support Vectors 来代替原有样本点集,用这组支持向量进行分类在结构风险最

小的意义下是最优的,通常支持向量数量远远小于原样本点的数量。由于高维空间中的线性函数集表达能力更强,容量更大,同时也意味着更容易发生过拟合。在支持向量机中,通过控制高维空间中的分类间隔来限制学习机器的容量,从而获得较强的推广能力。而核函数的使用,使得不必显式地进行非线性映射,只需在输入空间中进行运算,大大提高了效率。

因支持向量机的诸多优越性,文中采用了开源的 LibSVM 为实验工具进行了视像的镜头运动的检测。根据 LibSVM 对样本数据要求,把每一帧图像中宏块的运动矢量的 y 与 x 的比值以及运动矢量的模值作为

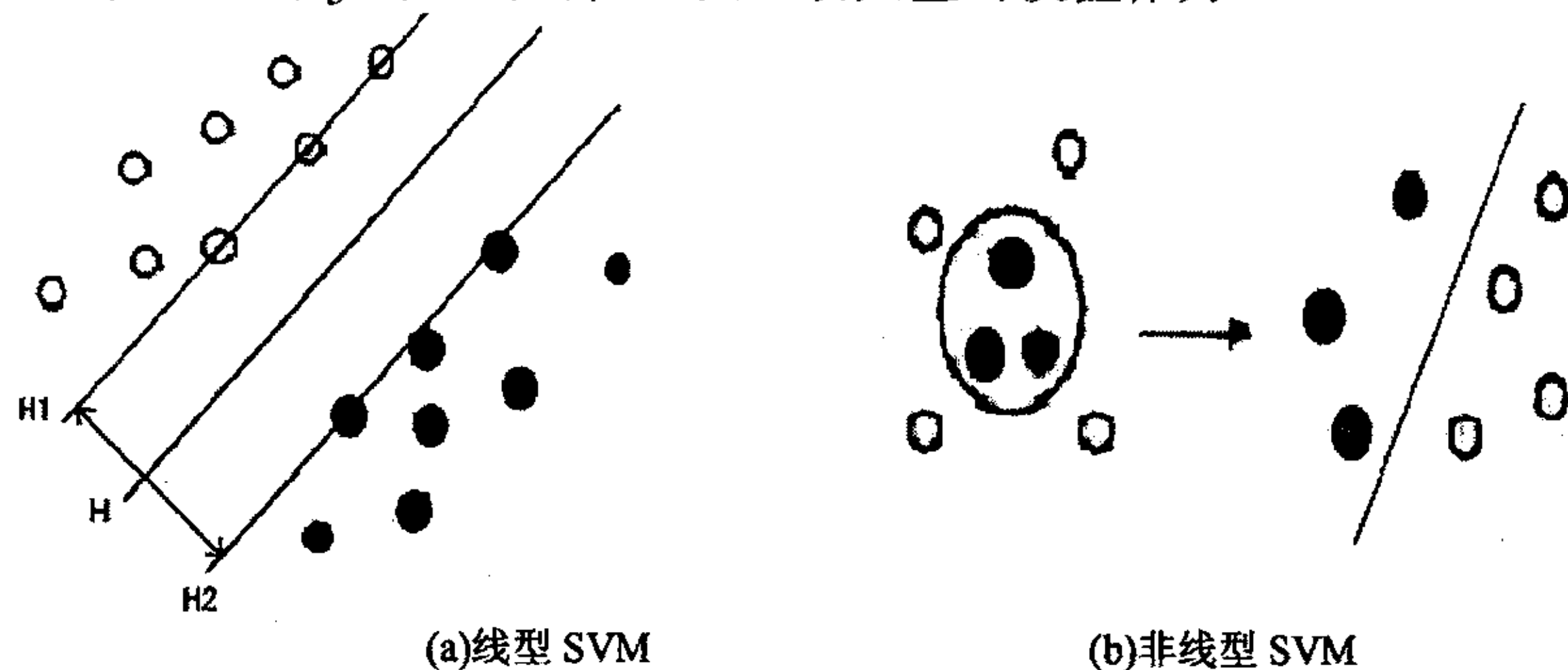


图2 基于支持向量机的二分类器示意

其样本的特征,这样每帧图像的特征维数就控制在 330×2 维,并把特征值较小的给以忽略,不参加支持向量机的训练和测试,从而达到减少实验的时间、加快实验的进程。由于摄像机的运动类型有三类,所以实验属于四分类的问题,为了对比实验的结果,我们还对样本特征进行了多次规划,分别对某种运动类型进行了二分类的 SVM 检测。实验的具体数据在下一小节中。

4 实验数据及结果

在第2小节中,讨论了采用了迭代的最小二乘法进行摄像机运动类型的检测流程,在本实验中,使用的是 TRECVID2005 中的部分数据,图3列举了其得到的准确率示意图。其横坐标代表了三组不同的实验参数,而纵坐标代表了实验的准确率。从这个图中,可以发现 PAN 的检索准确率最高,而 ZOOM 的准确率却较低。通过对数据的分析,可以发现在试验数据中,视像中摄像机的 PAN 运动远远多于 ZOOM 的运动类型,而且 PAN 运动的幅度也远远大于其他两种运动的幅度,通过迭代计算后得到的结果是 PAN 的准确率高其他两种运动。

图4是采用了支持向量机进行视像的摄像机运动检测的结果。前三个结果中的支持向量集中的核函数使用的是二次函数来完成的,后三组结果中使用的是

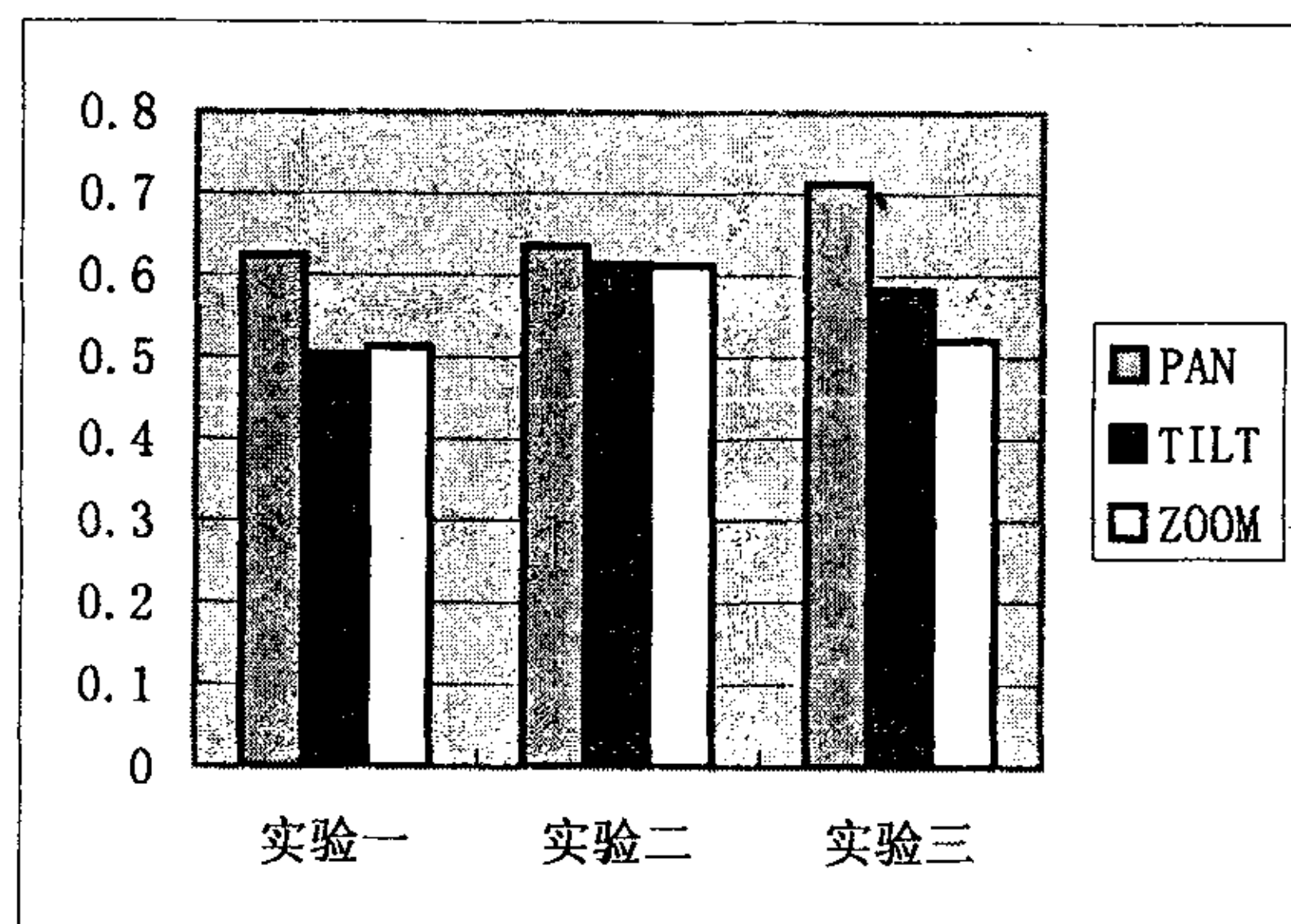


图3 采用最小二乘法的实验结果

多项式核函数来完成的实验。可以看出两者的数值相差不是非常的大,都比较准确地反映了视像中的摄像机的运动类型。最后一组实验是采用了四分类的 SVM 得出的实验结果,通过与前面的实验相比较,可以发现,在 LibSVM 众多分类与二分类问题都可以非常准确地给出其准确的答案,但是在实验的过程当中,可以明显地发现,采用四分类完成的实验比采用三个二分类实验要节省许多的时间。

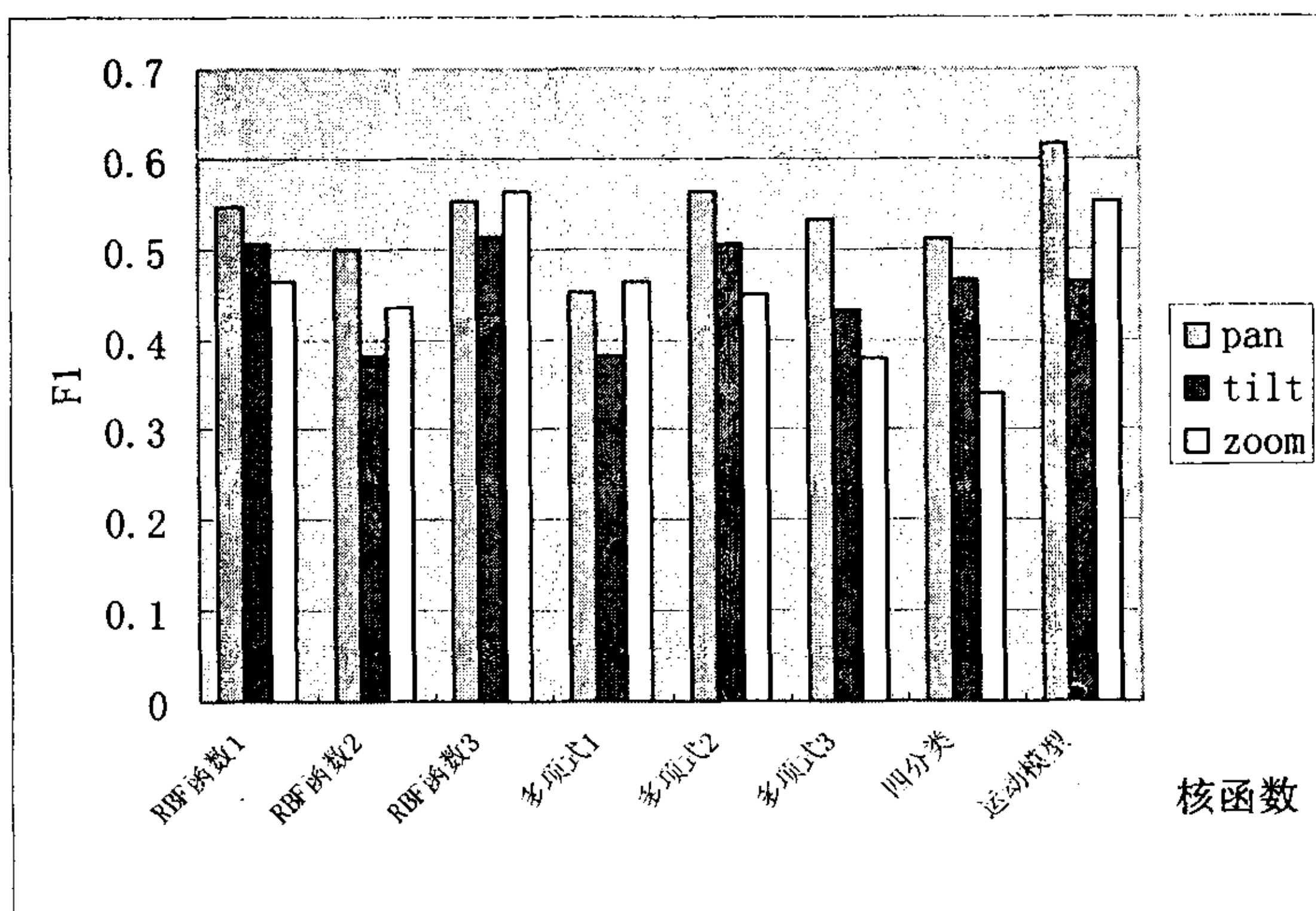


图4 采用 SVM 进行检测的实验结果

摄像机运动的检测是基于内容的视像检索的关键任务之一,在保证能准确检测出摄像机运动的情况下,对应的视像片段所代表的内容对人们进行视像的其他各种操作都会起到一定的作用。文中从两个不同的摄像机运动检测系统来分析和总结了运动的检测,因视像中的运动信息数据量大,处理的时间相对来说较长,特别是在用 SVM 进行分类检测的过程,因此对于视像的处理还有比较长的路要走。

参考文献:

- [1] Manjunath B S, Salembier P, Sikora T. Introduction to MPEG

式告知服务器其状态)

:State“* * * * *

:Sender <大气机> //Ar429:2,4//

:Receiver <服务器>

2.6 任务控制智能体

该智能体对任务序列进行分解并把任务分配给对应的执行者。对于一些复杂任务,智能体还将对执行者的能力和所需资源进行评估,确定由一个或多个执行者来承担。具体操作分以下二种情况:

(1) 查询任务分配表,若判定该任务由其它子系统承担,则以数据包形式通过高速通道(如:双口RAM/高速总线)把任务派发给相应的子系统。数据包中应包括任务的名称、优先级等,子系统执行结果通过高速通道回馈到核心计算机。

(2) 查询任务分配表,若任务需由多个执行者完成,则根据任务要求的执行时间评估该任务需由几个处理器来承担,若只需一个处理器,则在处理器群中指定一处理器,通过动态加载把任务加载到本地存储器中完成任务执行。若需多个处理器承担任务,则需通过“任务分配和通讯开销”均衡算法把任务分到不同的处理器上动态加载运行。

3 智能体数据交换机制

核心计算机智能体间使用共享内存方式进行数据交换,这块共享内存通常被称为“黑板”。每一智能体都把“黑板”作为自己的私有资源进行数据的存取和操作,使得同一个数据就能被所有智能体共享,从而达到数据交换的目的,当然这里要解决和避免内存操作冲突。显然,若“黑板”机制物理上采用非易失存储器(Non-Volatile)则可作为知识库和数据库的载体使用。核心计算机与外部智能体间的信息交换则通过“邮箱”方式进行。“邮箱”是一种双口RAM的机制,核心计算机每一智能体均有自己的“邮箱”。“信件”的分拣由核心计算机的通讯智能体负责。所以,核心计算机智能体可通过查询“黑板”和“邮箱”内容来获取内

部和外部的信息,当然也可通过它们来发送信息。

4 工作与展望

在核心计算机硬件架构建立以后,如何建立其高效和强有力的操作机制成为决定核心计算机智能水平和处理能力的关键。智能体的概念较好地表达了核心计算机的智能和并行特征,为核心计算机操作机制研究提供了理论基础,在智能体的理论框架下,文中仅对核心计算机的控制结构和各种功能进行了定义和描述,要建立核心计算机完善的操作机制有大量的开拓性工作要做,其中一些重要的基础性工作包括:

- (1)各智能体执行机的设计与验证;
- (2)智能体学习与成长机制的建立;
- (3)智能体间的任务协同与并行计算技术研究;
- (4)智能体动态任务重构技术研究;
- (5)数据库、知识库的构造与动态维护。

在完整的操作机制建立后,核心计算机可具备半自主、有限权限自主甚至全自主的运行能力^[5]。核心计算机智能和能力提升的途径主要通过神经网络的学习和成长、先进人工智能算法的更新、数据库和知识库等的扩充来完成,是逐步积累和完善的过程。当然,人工智能和智能计算机研究领域的每一点进展和突破都将为核心计算机的研究和发展带来新的曙光。

参考文献:

- [1] 陈国良,吴俊敏,章 锋.并行计算机体系结构[M].北京:高等教育出版社,2002.
- [2] Jordan H F. Fundamentals of Parallel Processing[M]. 北京:清华大学出版社,2003.
- [3] Meystel A M. Intelligent Systems: Architecture, Design and Control[M]. 北京:电子工业出版社,2003.
- [4] Hamid N, Earl B. Creating an adaptive embedded system by applying multi-agent techniques to reconfigurable hardware [J/OL]. 2004. www.sciencedirect.com.
- [5] Culler D E. Parallel Computer Architecture: A Hardware/Software Approach[M]. 北京:机械工业出版社,2003.
- [1] Bouthemmy P, Gelgon M, Ganansia F. A unified approach to shot change detection and camera motion characterization[J]. IEEE Transaction on Circuits and Systems for Video Technology, 1999, 9: 1030 - 1044.
- [2] 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 & Systems for Video Technology, 1999, 9: 1075 - 1099.
- [4] Tan Y P, Kulkarni S R, Ramadge P J. A new method for camera motion parameter estimation[C] // In Proceedings of International Conference Image Processing. [s. l.]: IEEE Computer Society, 1995: 405 - 408.
- [5] Boser B, Guyon I, Vapnik V. A training algorithm for optimal margin classifiers[C] // In: Haussler D, ed. Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory. New York, NY: ACM Press, 1992: 144 - 152.

(上接第 3 页)

- 7[M]. New York, NY, USA: John Wiley & Sons, Inc, 2002.