

基于轨迹行为模式特征的视频拷贝检测算法

郭俊波^{1,2)}, 李锦涛¹⁾, 张勇东¹⁾, 张冬明¹⁾, 吴 潇^{1,2)}

¹⁾(中国科学院计算技术研究所前瞻研究实验室 北京 100190)

²⁾(中国科学院研究生院 北京 100049)

(guojunbo@ict. ac. cn)

摘 要: 为了有效地利用视频的时域运动信息来提高视频拷贝检测的精度和鲁棒性,提出一种基于特征点轨迹行为模式的拷贝检测算法.首先从视频连续帧中提取特征点轨迹的行为模式特征,然后采用视觉关键词典技术构造视频的运动特征,最后基于运动特征的相似度进行视频拷贝检测.该算法在 TRECVID 标准数据集上取得了较高的检测精度.实验分析表明,基于轨迹的运动特征具有较强的描述区分能力,对各种常见的拷贝变化具有鲁棒性.

关键词: 视频拷贝检测;局部特征点检测;特征点轨迹;视觉关键词

中图分类号: TP391

Video Copy Detection Based on Trajectory Behavior Pattern

Guo Junbo^{1,2)}, Li Jintao¹⁾, Zhang Yongdong¹⁾, Zhang Dongming¹⁾, and Wu Xiao^{1,2)}

¹⁾(Center for Advanced Computing Research, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

²⁾(Graduate University of Chinese Academy of Sciences, Beijing 100049)

Abstract: In order to effectively extract temporal motion information of videos for robust and precise video copy detection, this paper proposes a novel video feature based on behavior patterns of keypoint trajectories. By the algorithm, first the keypoint trajectories are generated and modeled into behavior patterns. Then patterns are learned using bag-of-visual-word method as video features. Finally, similarity between features is used for video copy detection. Comparison experiments on benchmark video dataset (TRECVID) show that proposed algorithm achieves high detection precision. And this work demonstrates that the trajectory based feature has better robustness for various visual changes.

Key words: video copy detection; local feature point detection; keypoint trajectory; visual keywords

基于内容的视频拷贝检测技术(简称拷贝检测)利用视觉内容特征对视频进行比对,实现对特定视频内容的搜索、检测和跟踪,其在版权保护、视频数据管理、指定有害内容过滤等领域具有巨大的应用前景.视频拷贝检测的主要任务是在视频数据库中查找指定的查询视频片段,检测是否存在与查询视频片段内容相同的视频片段,如果存在,则称该视频片段为视频拷贝片段.

视频拷贝片段可能是源视频片段的直接复制,

也可能是对源视频经过各种变换(称为拷贝攻击)生成的,具体变换形式包括插入字幕或台标,片段组合,重新编码,画面图像的颜色、对比度、分辨率变化等.拷贝检测技术需要对各种拷贝攻击具有较好鲁棒性,其研究重点在于寻找有效的特征提取算法来表征视频的不变特性,进而提高拷贝检测对各种拷贝攻击的鲁棒性.本文研究主要基于视频行为特征的不变性,挖掘视频内容的不变视觉模式,并在此基础上进行视频拷贝检测.

收稿日期:2009-06-23;修回日期:2010-01-13.基金项目:国家“九七三”重点基础研究发展计划项目(2007CB311100);国家“八六三”高技术研究发展计划(2007AA01Z416);国家自然科学基金(60802028);北京市教育委员会共建项目专项.郭俊波(1975—),男,博士研究生,主要研究方向为多媒体检索;李锦涛(1962—),男,博士,研究员,博士生导师,主要研究方向为数字图像处理;张勇东(1973—),男,博士,研究员,博士生导师,主要研究方向为多媒体计算;张冬明(1977—),男,博士,副研究员,主要研究方向为多媒体计算;吴 潇(1982—),男,博士研究生,主要研究方向为视频内容分析.

1 相关工作

在视频拷贝检测技术领域,当前国内外研究重点是寻找有效的特征提取算法来提高拷贝检测对于各种拷贝攻击的鲁棒性.拷贝检测特征提取算法主要分为 2 类:基于视频帧图像的算法和基于视频时序信息的算法.

基于视频帧图像的特征提取算法^[1-6]在视频帧级别提取图像的全局或局部特征用于视觉内容匹配.文献[3]提出的签名特征是基于帧图像的全局统计特征的,如颜色直方图和边缘统计直方图.帧图像的全局特征对一些图像全局统一的变化(如全画面的模糊、偏移)具有鲁棒性.然而,图像局部变化(如字幕遮挡)会改变全局特征的分布,从而降低这类算法的匹配效果^[4].因此,文献[5-6]使用了基于特征点的图像局部特征应对较复杂的拷贝变化,以达到对图像的全局和局部变化都具有较好的鲁棒性.如文献[6]对每帧图像检测 Harris 特征点,提取特征点附近区域的边缘直方图作为局部特征.基于特征点的算法可以避免局部变化问题,但是由于特征点的数目巨大,造成对计算和存储的需求较大.另外,特征点特征维数低、信息量少,也使得特征点间的错配情况严重.

基于视频时序信息的特征提取算法^[7-9]利用视频序列的时序信息构造特征进行视频片段的配准.例如,基于图像分块有序序列的算法^[7-8]通过量化图像块间相对关系来构造不变视觉特征.文献[8]提出

将不同视频的有序序列用动态时间序列扭曲算法进行匹配.这类算法对全局变化情况下不对齐的帧图像序列匹配具有较好的鲁棒性,但是由于局部变化会打乱图像块间的有序关系,从而使得这类算法性能下降.同时该算法假设已知视频拷贝的帧序列的长度,使用滑动窗口的算法遍历式地进行序列匹配,运算量较大.

针对已有工作的不足,本文提出将相邻帧的特征点匹配连接为轨迹,利用轨迹丰富的行为特征来匹配视频片段,从而避免直接使用特征点特征进行匹配所造成的错配情况.近年来,视频中点或物体的轨迹作为一种相对稳定的视觉信息,已被应用于视频监控、物体跟踪等领域^[10-12].本文进一步利用视觉关键词技术,将视频镜头中提取的若干轨迹行为特征量化为统一的行为模式直方图用于视频匹配,也避免了以往局部特征匹配的大量计算消耗.

2 视频轨迹行为模式的提取

本文提出一种利用视频行为特征的拷贝检测算法,该算法将视频拷贝检测过程分成 2 个阶段:离线视频库索引阶段负责对视频库中的视频片段进行特征提取并构建视频库模式词典;在线匹配阶段对查询视频提取特征并与视频模式进行比对,进而发现视频拷贝.

本文算法框架如图 1 所示.首先离线计算视频库模式词典,对视频库中的每个视频提取其特征点轨迹,并对轨迹行为模式进行量化编码成 5 维特征

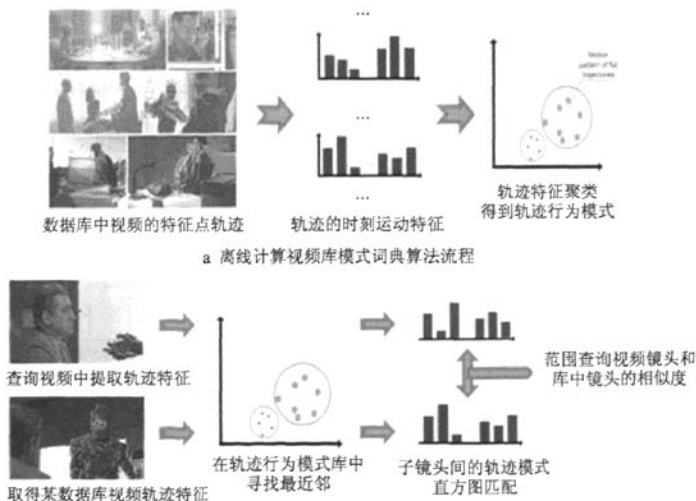


图 1 本文算法流程示意图

向量;然后基于 K -means 聚类算法生成轨迹行为模式词典,并将每个镜头表示为轨迹词频向量,如图 1 a 所示.由于特征点轨迹对各种拷贝变化具有鲁棒性,所生成的词频向量具有良好的代表性,可用于计算视频镜头间的相似度.在在线匹配(拷贝检测)阶段,首先将查询视频镜头表示为词频向量,然后与视频库中镜头的轨迹词频向量进行匹配,从而判断查询视频是否为视频拷贝,如图 1 b 所示.

特征点检测和跟踪是提取连续帧轨迹的两大技术要点.本文提出一种融合 Harris 特征点检测与 KLT 点跟踪算法的快速特征点轨迹提取算法.Harris 特征点检测是一种公认的对各种拷贝攻击比较鲁棒的局部特征提取算法^[5],而 KLT 算法则是一种快速的局部特征点跟踪算法.本文算法轨迹提取过程为,首先利用 Harris 算子检测首个视频帧的所有特征点,然后使用 KLT 算法在后续帧快速地跟踪这些特征点.除 Harris 外,常用的特征点检测算法还包括 Hessian-Affine 和 Difference-of-Gaussian (DoG)^[5]等.文献[13]采用 DoG 检测特征点并基于点的 SIFT 特征进行匹配提取轨迹,将其用于识别人体动作,基于 SIFT 特征的轨迹提取算法需要对每个视频帧检测特征点,导致计算量巨大,且 SIFT 特征不具备时域信息.本文采用的点跟踪算法只需要在少量帧上检测特征点,更适合拷贝检测应用.

帧图像的 Harris 特征点检测算法为

$$M = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (1)$$

$$R = \det(M) - k \times \text{tr}(M)^2, k \in [0, 0.04] \quad (2)$$

式(1)为 Harris 算子使用的自相关矩阵,其中, I 为图像像素矩阵, (x, y) 为相应图像像素坐标, I_x , I_y 分别为图像像素在水平方向和垂直方向的一阶偏微分, I_x^2 和 I_y^2 分别为图像像素在水平方向和垂直方向的二阶偏微分.响应函数式(2)用于计算图像像素的响应值 R ,选取 R 值排序前 N 个像素点作为 Harris 特征点.基于 M 矩阵的 R 值物理意义是,表征点所在小图像区域的变化稳定性,通常图像边缘和物体棱角上的像素满足变化鲁棒性要求,被检测为特征点. Harris 特征点在复杂自然图像上有着良好的性能,特别是在旋转和亮度变化等条件下依然保持较高的可重复率;同时, Harris 算法检测过程计算复杂度比其他特征点检测算法低^[5],适合视频应用.

在视频第一帧上检测到 N 个 Harris 特征点(实验中取 $N = 1000$)后,从第二视频帧开始使用 KLT 算法跟踪所有特征点. KLT 算法是一种稀疏

光流跟踪算法^[14],其跟踪过程如下:通过求特征点 p 在 x 和 y 方向上的速度,来计算该点在下一帧出现的位置,从而实现跟踪.

另外,在点跟踪过程中,当视频内容发生剧烈变化,如镜头场景切换、视频 2 个相邻帧内容完全不一致时, KLT 算法无法跟踪轨迹,造成轨迹断开.本文算法设定当轨迹条数突然下降,即少于一定阈值(实验中取 30)时,重新检测下一帧的 Harris 特征点并重新使用 KLT 算法跟踪,如图 2 所示形成新的轨迹簇.本文把轨迹簇断开处作为边界,称边界之间的视频子段为亚镜头,以亚镜头为单位提取轨迹行为模式特征并进行拷贝检测.

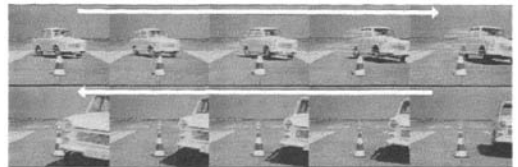


图 2 使用 Harris 和 KLT 算法产生的视频轨迹

视频轨迹包含了帧图像局部特征点在连续帧的运动行为信息,这些特征点大部分能在各种拷贝变化下存在,从而保证拷贝检测的精度.然而,轨迹的长度和数量都随着视频内容变化而变化,为了取得统一的轨迹描述,本文提出一种轨迹行为编码算法来建模轨迹行为.如图 3 所示,轨迹上的特征点相对位置的变化被编码为 5 种代码,一条轨迹可以被表示一个编码序列.具体而言,首先记录某特征点在连续帧上的坐标位置;然后根据 4 个象限加中心区域的编码模板,将轨迹点的前后相对位置建模为行为特征,其中中心区域的大小定义了静态特征点的运动范围,设定为较小的 7×7 像素区域;最后将一条轨迹转变为一组编码数字.

单个代码存储了轨迹的空间信息,代码的顺序表征了轨迹的时间信息.由于 Harris 特征点检测算法和 KLT 算法的鲁棒性,常见的全局拷贝变化一般

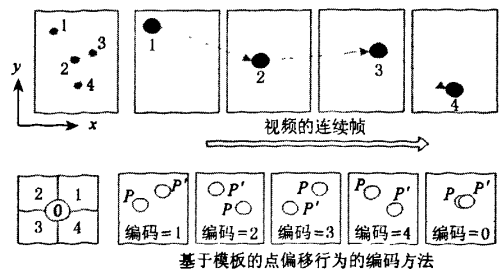


图 3 轨迹行为模式特征的量化过程

只会轻微影响轨迹的数目和长度.但局部变化(如遮挡等)和时域的拷贝变化(如丢失帧、帧率降低等)会改变轨迹的数目和亚镜头分割的位置(即改变了轨迹的长度),进而改变其行为模式.针对这些问题,本文提出一种简单的特征提取算法,通过对轨迹行为编码数列进行归一化统计,生成 5 维直方图,平滑掉各种变化带来的影响,即

$$CH(i) = \frac{\#i}{\text{Length}(\text{Code_Sequence})}$$

例如,序列{4,1,3}被转换的直方图为{0:0,1:0.33,2:0,3:0.33,4:0.33}.轨迹越长,直方图的区别性越强.本文使用这种简单的转换保留轨迹行为的统计信息,消除了大部分各种变化可能带来的噪声影响,并方便轨迹之间的比较和轨迹聚类.

3 视频轨迹行为模式的匹配

经过提取视频不变视觉模式(即轨迹行为模式),一个视频被分割为若干的亚镜头;每个亚镜头都包含一组轨迹特征直方图.类似于图像检索中使用的视觉关键词技术^[15],本文把亚镜头看作包含一组轨迹关键词的文档.下面基于这类视觉关键词的思想提出不变视觉模式匹配算法,用于拷贝检测.

3.1 生成轨迹关键词典

利用视觉关键词技术,在查询之前对源视频库中的所有视频提取轨迹,并将其建模为大量的 5 维直方图特征;然后利用经典 K-means 算法进行聚类,产生一个轨迹行为特征词典,每个关键词都是一个有代表性的 5 维特征. K-means 把大量轨迹行为进行聚类,去除了数据冗余并削弱了噪声的影响.假设生成 K 个聚类中心作为关键词,每个亚镜头 s 可以表示为 K 维词频向量 $V_s = \{t_1, t_2, \dots, t_i, \dots, t_k\}$, t_i 表示轨迹特征关键词 i 在亚镜头 s 中出现的次数. K-means 算法得到的词典的规模 K 必须人为指定,

K 的设置对聚类效果有很大影响.对于本文中使用的数据集,我们通过测试不同 K 值下的检测精度,得到一个最优 K 值.

在实验中也观察到,静态镜头中大量的特征点保持位置不变,其特征数列为全 0 序列,一些短轨迹也缺乏区别性,这类在文本检索和图像检索中共有的现象可称为停用词现象.参考文本检索的经典思想,本文利用 tf-idf 公式

$$t_i = \frac{n_{id}}{n_d} \log \frac{N}{n_i}$$

对亚镜头词频向量做加权操作.其中, n_{id} 是关键词 i 在亚镜头 d 中出现次数, n_d 是亚镜头 d 包含的关键词个数, n_i 表示关键词 i 在整个视频库所有亚镜头中出现的次数, N 表示视频库中亚镜头的总数.使用 tf-idf 加权后,由于停用词 n_i 数目大,自动降低了其权值和影响力,达到了去停用词的目的.

3.2 基于亚镜头的拷贝匹配和定位

在线查询过程中,首先对查询视频提取轨迹并将其分割为若干个亚镜头,然后计算每个亚镜头的 5 维特征并由关键词词典归类为关键词,每个查询亚镜头被表示为词频向量.

查询亚镜头和源视频库中的亚镜头的相似度的计算可以使用多种算法.本文使用余弦距离来计算亚镜头 S_a, S_b 间的相似度

$$\text{sim}(S_a, S_b) = 1 / \left(1 - \frac{V_a \times V_b'}{(V_a' \times V_a)^{\frac{1}{2}} (V_b' \times V_b)^{\frac{1}{2}}} \right),$$

即使用 2 个向量 V_a 和 V_b 的夹角余弦值来度量其相似性.全部亚镜头的两两相似度组成相似度矩阵,如图 4 a 所示.

拷贝检测的目的是要定位查询视频与源视频共有片段的起止位置.基于亚镜头相似度矩阵,除了正确的匹配外,还可能存在零散的错匹配.为了去除这

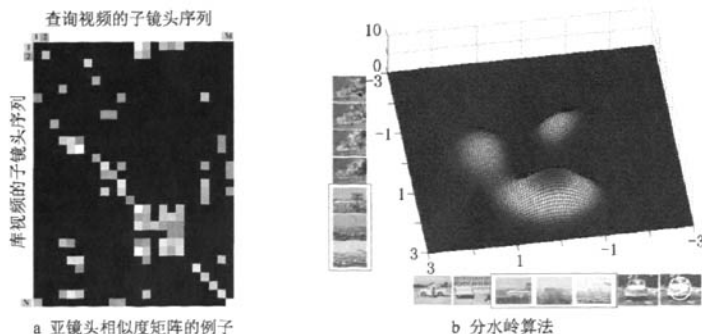


图 4 基于分水岭算法的相似镜头段落检测

些噪声,将此矩阵作为图像 I ,对 I 使用经典的分水岭算法进行分割.如图 4b 所示,在实验中,正确匹配段常常表现为连续型的“山脊”或“高原”的图像块状局部结构,计算这些块结构的相似度之和 $\sum sim_{ij}$,使用一个阈值 θ 过滤掉高频错配噪声,从而发现正确匹配的位置.通常选择其 $\sum sim_{ij}$ 最大的部分作为拷贝片段,实验中,我们经验性地选择为 $\theta=0.5$.

4 实验及结果分析

本文实验的源视频库采用TRECVID 2008年拷

贝检测比赛^①使用的数据集,查询视频的制作遵循该拷贝检测比赛使用的算法.我们从 20 个源视频中截取小段(3 s~60 s)嵌入到 20 个短视频中作为拷贝查询,另外选取 10 个短视频作为非拷贝查询,对这 30 个查询视频分别施加该数据集中定义的拷贝变化.

如第 2 节所述,适合拷贝检测的视觉特征必须满足鲁棒性、实时性要求.表 1 所示为本文实验数据处理的具体情况,采用本文提出的拷贝检测算法在特征提取上能达到实时速度(即处理时间等于或者短于视频播放时间),并占用相对少量的存储空间,适合大规模的视频拷贝检测任务.

表 1 实验数据集的相关信息

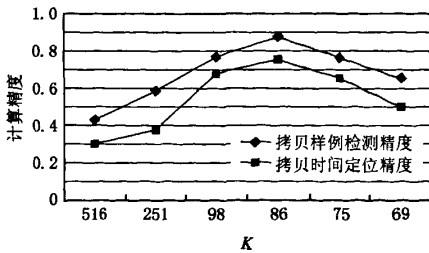
视频集合	视频个数	分割出的亚镜头个数	播放时间/h	系统处理耗时/h	视频文件占用存储/GB	特征文件占用存储/MB
源视频	101	652500	58	25	100	227.69
查询视频	30	25500	2.33	0.93	5	7.22

首先,使用 TRECVID 2008 年拷贝检测比赛的评测指标来验证本文算法的鲁棒性.评测指标包括检测准确率

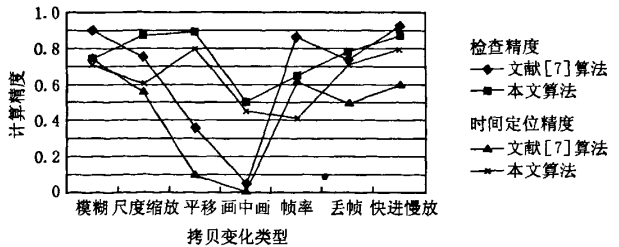
$$Det_Prec = \frac{\#Correct_Det}{\#All_Det}$$

和定位精度

$$Loc_Accu = \frac{\#Overlap(Det, Copy)}{\#Copy_Frames}$$



a 不同 K 值下的平均检测性能



b 各种拷贝变化下与文献[7]算法的对比

图 5 本文算法和已有算法在标准数据集上的检测精度

其次对比了本文算法和其他算法的性能. OM 算法是常见的一种拷贝检测算法^[3,7-8], Chiu 等^[7]算法声称比传统算法效果更好.当 $K=86$ 时,我们使用一些比较难处理的拷贝变化下的查询视频,分别使用本文算法和文献^[7]算法进行检测,得到的评测指标曲线如图 5b 所示.图 5b 表明,本文算法使用局部稳定特征点的轨迹行为特征,能够对一些比较复杂的局部变化(如遮挡、裁剪画面、画面平移和画

中画等)具有更好的鲁棒性.这些局部图像变化改变了帧图像的部分内容,打乱了 OM 算法依赖的块序列关系,从而影响了其鲁棒性.我们针对视频时空不变模式进行建模,把拷贝变化前后不变的视觉内容作为特征进行匹配,所以具有鲁棒性.但是,图 5b 也表明,本文算法对一些时域变化并不能取得更好

① <http://www-nlpir.nist.gov/projects/tv2008/tv2008.html>

的效果,其原因是当前使用的简单行为统计算法和 5 维直方图特征不能完全平滑地去掉时序变化的影响。

在计算效率上,本文算法比文献[7]算法有较大提升。文献[7]算法中使用动态时间扭曲对齐进行拷贝匹配,属于动态规划算法,其复杂度为 $O(n^2)$, n 为视频帧数。而本文算法基于亚镜头进行匹配,理论

复杂度为 $O(m^2)$, m 为视频镜头数目,在 $m \ll n$ 的情况下可以节省大量计算。轨迹提取过程采用 OpenCV 函数,可达到 90 帧/s 的处理速度。为了体现局部特征轨迹的效果,我们在图 6 中给出了在各种拷贝情况下亚镜头轨迹的形态,表明在拷贝变化后,轨迹行为的总体趋势相似。图 6 中一幅图像是一个亚镜头的代表帧,其中绿色曲线表示亚镜头的特征点轨迹。



图 6 在各种常见拷贝变化下轨迹行为的一致性

5 总 结

如何提取鲁棒的视觉特征是基于内容的视频拷贝检测的一个关键研究点。本文算法基于视频连续帧的轨迹行为,提出一种精简的不变模式特征,在计算时具有高效性;并使用轨迹视觉关键词技术,在亚镜头级别定位拷贝视频片段。在标准数据集上的对比实验表明,与已有算法相比,本文算法对各种变化,尤其是局部画面变化有更好的鲁棒性。

参考文献 (References):

- [1] Joly A, Buisson O, Frélicot C. Content-based copy retrieval using distortion-based probabilistic similarity search [J]. *IEEE Transactions on Multimedia*, 2007, 9(2): 293-306
- [2] Wu X, Hauptmann A G, Ngo C W. Practical elimination of near-duplicates from web video search [C] // *Proceedings of ACM International Conference on Multimedia*, Augsburg, 2007: 218-227
- [3] Kim C, Vasudev B. Spatiotemporal sequence matching for efficient video copy detection [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2005, 15(1): 127-132
- [4] Gao Ke, Lin Shouxun, Zhang Yongdong, et al. Object-based image retrieval using spatial context [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2008, 20(11): 1452-1458 (in Chinese)
(高科, 林守勋, 张勇东, 等. 基于空间上下文的目标图像检索[J]. *计算机辅助设计与图形学学报*, 2008, 20(11): 1452-1458)
- [5] Mikolajczyk K, Tuytelaars T, Schmid C, et al. A comparison of affine region detectors [J]. *International Journal of Computer Vision*, 2005, 65(1/2): 43-72
- [6] Law-to J, Gouet-brunet V, Buisson O, et al. Local behaviours labelling for content based video copy detection [C] // *Proceedings of the 18th International Conference on Pattern Recognition*, Hong Kong, 2006, 232-235
- [7] Chiu C Y, Li C H, Wang H A, et al. A time warping based approach for video copy detection [C] // *Proceedings of the 18th International Conference on Pattern Recognition*, Hong Kong, 2006: 228-231
- [8] Hua X S, Chen X, Zhang H J. Robust video signature based on ordinal measure [C] // *Proceedings of International Conference on Image Processing*, Singapore, 2004: 685-688
- [9] Hampapur A, Hyun K H, Bolle R. Comparison of sequence matching techniques for video copy detection [C] // *Proceedings of the SPIE*, San Jose, 2002, 4676: 194-201
- [10] Law-to J, Chen L, Joly A, et al. Video copy detection: a comparative study [C] // *Proceedings of Conference on Image and Video Retrieval*, Amsterdam, 2007: 371-378

(下转第 958 页)

- (王 琨, 郑南宁. 基于 SFM 算法的三维人脸模型重建[J]. 计算机学报, 2005, 28(6): 1048-1053)
- [7] Su Congyong, Zhuang Yueting, Huang Li, *et al.* Analysis-by-synthesis approach for facial modeling based on orthogonal images [J]. *Journal of Zhejiang University: Engineering Science*, 2005, 39(2): 175-179 (in Chinese)
(苏从勇, 庄越挺, 黄 丽, 等. 基于正交图像生成人脸模型的合成分析方法[J]. 浙江大学学报: 工学版, 2005, 39(2): 175-179)
- [8] Zhang L, Snavely N, Curless B, *et al.* Spacetime faces: high resolution capture for modeling and animation [J]. *ACM Transactions on Graphics*, 2004, 23(3): 548-558
- [9] Amin S H, Gillies D. Analysis of 3D face reconstruction [C] // *Proceedings of the 14th International Conference on Image Analysis and Processing, Modena*, 2007: 413-418
- [10] Fidaleo D, Medioni G. Model-assisted 3D face reconstruction from video [M] // *Lecture Notes in Computer Science*. Heidelberg: Springer, 2007, 4778: 124-138
- [11] Liu Z C, Zhang Z Y, Jacobs C, *et al.* Rapid modeling of animated faces from video images [C] // *Proceedings of ACM Multimedia, Los Angeles*, 2000: 475-476
- [12] Cootes T F, Edwards G, Taylor C J. Comparing active shape models with active appearance models [C] // *Proceedings of British Machine Vision Conference, Nottingham*, 1999: 173-182
- [13] Baker S, Matthews I. Lucas-Kanade 20 years on, a unifying framework [J]. *International Journal of Computer Vision*, 2004, 56(3): 221-255
- [14] Bregler C, Hertzmann A, Biermann H. Recovering non-rigid 3D shape from image streams [C] // *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island*, 2000: 690-696
- [15] Xiao J, Kanade T. Non-rigid shape and motion recovery: degenerate deformations [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Washington D C, 2004: 668-675
- [16] Turk M, Pentland A. Eigenfaces for recognition [J]. *Journal of Cognitive Neuroscience*, 1991, 3(1): 71-86

~~~~~

(上接第 948 页)

- [11] Möenne-loccoz N, Bruno E, Marchand-Maillet S. Local feature trajectories for efficient event-based indexing of video sequences [M] // *Lecture Notes in Computer Science*. Heidelberg: Springer, 2006, 4071: 82-91
- [12] Satoh S, Takimoto M, Adachi J. Scene duplicate detection from videos based on trajectories of feature points [C] // *Proceedings of Conference on Multimedia Information Retrieval, Augsburg*, 2007: 237-244
- [13] Sun J, Wu X, Yan S C, *et al.* Hierarchical spatio-temporal context modeling for action recognition [C] // *Proceedings of International Conference on Computer Vision and Pattern Recognition, Miami*, 2009: 2004-2011
- [14] Shi J B, Tomasi C. Good features to track [C] // *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, Seattle*, 1994: 593-600
- [15] Sivic J, Zisserman A. Video Google: a text retrieval approach to object matching in videos [C] // *Proceedings of International Conference on Computer Vision, Nice*, 2003, 2: 1470-1477

# 基于轨迹行为模式特征的视频拷贝检测算法

作者: 郭俊波, 李锦涛, 张勇东, 张冬明, 吴潇, Guo Junbo, Li Jintao, Zhang Yongdong, Zhang Dongming, Wu Xiao  
作者单位: 郭俊波, 吴潇, Guo Junbo, Wu Xiao (中国科学院计算技术研究所前瞻研究实验室, 北京, 100190; 中国科学院研究生院, 北京, 100049), 李锦涛, 张勇东, 张冬明, Li Jintao, Zhang Yongdong, Zhang Dongming (中国科学院计算技术研究所前瞻研究实验室, 北京, 100190)  
刊名: 计算机辅助设计与图形学学报 **ISTIC EI PKU**  
英文刊名: JOURNAL OF COMPUTER-AIDED DESIGN & COMPUTER GRAPHICS  
年, 卷(期): 2010, 22(6)  
被引用次数: 0次

## 参考文献(15条)

1. Joly A, Buisson O, Frélicot C Content-based copy retrieval using distortion-based probabilistic similarity search 2007(2)
2. Wu X, Hauptmann A G, Ngo C W Practical elimination of near-duplicates from web video search 2007
3. Kim C, Vasudev B Spatiotemporal sequence matching for efficient video copy detection 2005(1)
4. 高科, 林守勋, 张勇东, 唐胜 基于空间上下文的目标图像检索[期刊论文]-计算机辅助设计与图形学学报 2008(11)
5. Mikolajczyk K, Tuytelaars T, Schmid C A comparison of affine region detectors 2005(12)
6. Law-to J, Gouet-brunet V, Buisson O Local behaviours labelling for content based video copy detection 2006
7. Chiu C Y, Li C H, Wang H A A time warping based approach for video copy detection 2006
8. Hua X S, Chen X, Zhang H J Robust video signature based on ordinal measure 2004
9. Hampapur A, Hyun K H, Bolle R Comparison of sequence matching techniques for video copy detection 2002
10. Law-to J, Chen L, Joly A Video copy detection: a comparative study 2007
11. M(o)enne-locco N, Bruno E, Marchand-Maillet S Local feature trajectories for efficient event-based indexing of video sequences 2006
12. Satoh S, Takimoto M, Adachi J Scene duplicate detection from videos based on trajectories of feature points 2007
13. Sun J, Wu X, Yan S C Hierarchical spatio-temporal-context modeling for action recognition 2009
14. Shi J B, Tomasi C Good features to track 1994
15. Sivic J, Zisserman A Video Google: a text retrieval approach to object matching in videos 2003

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_jsjfdsjytxxb201006006.aspx](http://d.g.wanfangdata.com.cn/Periodical_jsjfdsjytxxb201006006.aspx)

授权使用: 中科院计算所(zkyjsc), 授权号: fca16639-9031-4362-b2cb-9e4e0127f3c0

下载时间: 2010年12月16日