

# 基于时序特征的视频相似性匹配算法

贾克斌, 邓智玘, 庄新月

(北京工业大学 电子信息与控制学院, 北京 100022)

**摘要:** 针对视频亮度整体漂移、突变干扰、视频再编辑引起的帧率变化问题, 提出了采用视频时序特征来描述视频内容, 描述了视频的时间、空间信息. 采用梯度匹配算法和改进的动态时间规划算法进行相似性匹配. 实验采用了80段广告视频和200段再编辑视频对算法进行检测, 结果证明了算法的有效性和鲁棒性.

**关键词:** 视频信号处理; 基于内容的检索; 查询视频; 时序特征; 梯度方法; 动态规划

**中图分类号:** TN 911.73

**文献标识码:** A

**文章编号:** 0254-0037(2008)12-1250-04

用户如何在浩如烟海的信息中快速获得自己需要的内容, 使得基于内容的视频检索<sup>[1-2]</sup>方式对视频信息进行浏览、查询的要求越来越普遍.

对视频内容的已有描述大多借助于视频的静态特征(基于图像的特征)、运动特征和关键帧等技术<sup>[3-5]</sup>, 它们都存在着各种缺陷, 或在应用范围上受到一定的限制. MPEG-7视觉描述引入了视频时序特征<sup>[6]</sup>的概念, 该方法能有效地运算, 但结果容易受到视频噪音的干扰. 本文提出视频时序特征, 实验证明视频时序特征能有效的匹配视频序列, 同时, 引入了异常因子来解决视频传播过程中遇到的亮度整体漂移和突变强干扰问题, 并提出了一种改进的动态时间规划(DTW)算法来处理电视信道干扰(如颜色突变)和视频再编辑(快慢镜头)中遇到的问题.

## 1 视频时序特征的提取

提取视频时序特征时需要确定颜色空间和分割子区域的个数. 由于人眼对亮度信号比较敏感, 在实际系统中, 逐帧提取静态图像序列的平均亮度值作为视频时序特征. 设视频流  $D = \{F_1, F_2, F_3, \dots, F_N\}$ , 用  $Y_k$  表示视频流第  $k$  帧图像的平均亮度值, 得到视频时序特征曲线  $S = \{Y_1, Y_2, \dots, Y_N\}$ . 进一步, 为了提高检索的精确度, 增加视频时序特征中含有的信息量, 将视频中每帧静态图像分割成几个子区域, 若一幅图像具有  $M \times N$  个像素, 则极限情况下, 图像可被分为  $M \times N$  个区域, 如图1所示.

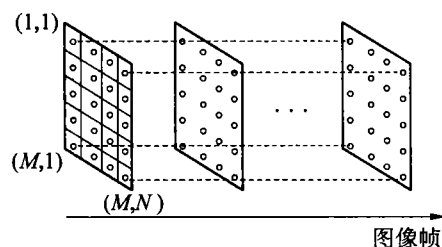


图1 图像分割方法

Fig.1 Frame divided method

## 2 相似性匹配算法

### 2.1 直接比对匹配算法

由视频时序特征的定义可知, 内容相同的视频具有相同的视频时序特征, 直接比对匹配算法逐帧累计2段视频的亮度特征差值, 再与阈值  $L$  比较, 判断是否在容忍范围内. 基本原理如图2所示. 设  $m$  帧样例

收稿日期: 2007-06-20.

基金项目: 国家自然科学基金项目资助(60672050, 60431020).

作者简介: 贾克斌(1962-), 男, 河南安阳人, 教授.

视频  $p$  与  $n$  帧目标视频  $q$  ( $m < n$ ), 算法分为 3 步.

第 1 步 对样例视频和目标视频逐帧计算平均亮度  $Y_{pi}$  和  $Y_{qi}$ , 得到视频时序特征  $S_p$  和  $S_q$ .

第 2 步 将样例视频时序特征第 1 帧对应目标视频时序特征的第 1 帧, 逐帧累计两视频时序特征的对应帧差绝对值, 得到帧差之和  $T_1$ .

第 3 步 若  $T_1 < L$ , 则认为匹配成功, 否则样例视频的第 1 帧移动到对应目标视频的第 2 帧开始比较, 重复第 2 步的比较过程得到  $T_2$ , 若  $T_2 > L$ , 继续移动样例视频时序特征, 直到  $T_j < L$ , 匹配成功, 或者当  $j > n$  时, 匹配失败.

$$T_j = \sum_{i=1}^m (|Y_{q(j+i+1)} - Y_{pi}|) \quad (j=1, \dots, n) \quad (1)$$

## 2.2 梯度匹配算法

梯度匹配算法, 将视频时序特征曲线的相似性匹配问题转化为二元函数特征曲线匹配的问题. 将曲线的变化趋势作为曲线匹配的特征, 并增加异常因子以去除少量强干扰的影响, 具体算法分 4 步.

第 1 步 设置每帧图像的分割数为  $L \times V$ , 并逐帧计算样例视频  $p$  和目标视频  $q$  当中  $L \times V$  个子区域里的平均亮度值, 得到视频时序特征曲线  $S_p$  和  $S_q$ .

$$\Delta Y_{pf} = |Y_{p(f+1)} - Y_{pf}| \quad f=1, \dots, m-1 \quad (2)$$

式(2)中,  $\Delta Y_{pf}$  为样例视频前后帧的平均亮度差;  $f$  为样例视频帧号.

$$\Delta Y_{qg} = |Y_{q(g+1)} - Y_{qg}| \quad g=1, \dots, n-1 \quad (3)$$

式(3)中,  $\Delta Y_{qg}$  为目标视频前后帧的平均亮度差;  $g$  为目标视频帧号.

第 2 步 若  $g$  小于  $n - m/2$  且  $f$  小于  $n$ , 计算样例视频和目标视频各个子区域时序特征的变化值  $\Delta Y_p$  与  $\Delta Y_q$  的差, 取其绝对值

$$\Delta T_f = |\Delta Y_{q(g+f)} - \Delta Y_{pf}| \quad (4)$$

第 3 步 若  $\Delta T_f < M$  ( $M = L/m$ ), 则认为样例视频第  $f$  帧与目标视频第  $(g+f)$  帧的对应子区域匹配, 则可接受的帧数加 1; 在进行完一次样例视频和目标视频对应子区域循环比较后, 计算异常因子值. 若异常因子小于用户设定的阈值, 认为 2 段视频匹配成功

$$\text{异常因子} = \left( 1 - \frac{\text{可接受的帧数}}{\text{样例视频总帧数}} \times 100\% \right) \quad (5)$$

第 4 步 否则,  $g = g + 1$ , 移动样例视频曲线重复第 2 步的比较过程, 直至循环结束.

## 2.3 改进的动态时间规划算法

在实际应用中, 视频节目还存在着帧率变化和视频再编辑(如快慢动作)问题. 为此, 提出了一种改进的动态时间规划(DTW)算法. 传统的动态时间规划算法<sup>[7-8]</sup>有对称算法和非对称算法两类. 具体来说, 对 2 个视频片段 A 和 B, 假定分别包含  $I$  帧和  $J$  帧. 可以首先计算 A 序列和 B 序列中任意 2 个帧对的差值并预先储存, 得到  $I \times J$  的差值矩阵

$$D(i, j) = d(i, j) + \min(D(i-1, j), D(i-1, j-1), D(i, j-1)) \quad (6)$$

式中  $d(i, j)$  为第  $i$  行第  $j$  列的元素, 对称算法通过循环计算差值来得到最优比对结果.

非对称动态时间规划算法

$$D(i, j) = d(i, j) + \min(D(i-1, j-1), D(i-1, j-2)) \quad (7)$$

考虑到对视频序列的比对过程中, 待比对的视频同时包括快动作和慢动作, 允许跳帧和重复帧. 本文采用对称算法和非对称算法合并的方案

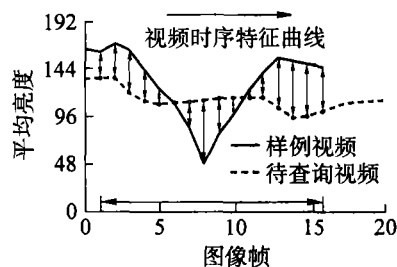


图 2 直接比对匹配算法原理

Fig. 2 The principle of direct comparing algorithm

$$D(i, j) = d(i, j) + \min(D(i-1, j-1), D(i-1, j-2), D(i, j-1)) \quad (8)$$

为了避免出现错误的检测结果,可以反向跟踪得到对应最终结果的匹配路径.考虑到视频编辑过程所对应的视频帧率变化为固定比率.如果匹配结果对应的视频片段是通过视频编辑得到的,则对应的路径所代表的一组点应当近似呈现一条直线.通过检测路径是否接近直线,可以判断是否是真正的匹配序列.算法回溯路径如图3所示,每个交叉点表示对应的样例视频与待查询视频相匹配的图像帧帧号,各个交叉点连起来形成一条起始于(1,1)而终止于(I,J)的路径.

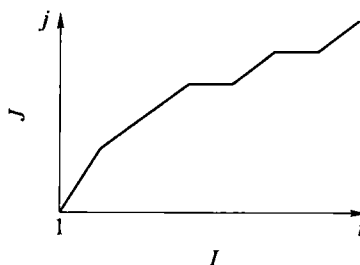


图3 动态时间规划算法回溯路径

Fig.3 The trace path of DTW algorithm

### 3 实验结果及分析

实验一 采用80段视频来测试梯度匹配算法.为便于分析,按干扰类型将测试视频分为3类:1)存在亮度整体漂移;2)存在突变干扰;3)存在轻微干扰.实验数据集中第1种情况28段视频,第2种情况16段视频,第3种情况36段视频.一段广告在不同时段采集2~5组.任选一段视频作为样例视频进行检索.根据干扰强度,适当调整参数.每帧分为 $2 \times 2$ 子区域,阈值为4~6,异常因子取1%~5%.实验结果如表1所示.其中,查全率 =  $\frac{\text{有关联的正确检索结果}}{\text{所有有关系的结果}}$ ;准确度 =  $\frac{\text{有关联的正确检索结果}}{\text{所有检索到的结果}}$

表1 不同干扰类型的效果评价

Table 1 Experimental results according to different kind of disturbance

干扰类型	检索类型	正确检索	误检	漏检	查全率/%	准确度/%
1	直接法	19	0	11	63.33	100
	梯度法	29	17	1	96.66	63.04
2	直接法	12	0	4	75	100
	梯度法	16	0	0	100	100
3	直接法	30	0	4	88.24	100
	梯度法	34	11	0	100	75.56

原有直接比对匹配算法不能处理存在曲线整体飘移干扰的情况,查全率仅63.33%.而改进算法将查全率提高到96.66%.对于另外2种干扰情况,改进算法均能达到较好的效果.改进方法在准确度方面虽然有一定程度的降低,但影响不大.改进算法在相似性匹配部分耗时是原算法的1.5倍,但速度仍可接受,每1000帧匹配耗时在 $10^{-3}$ ms级.采用隔行扫描后,平均耗时从5s/100帧提高到4s/100帧,速度提高了20%.

实验二 为了验证改进的动态时间规划算法的有效性,选择了200个视频广告做为实验数据集合.数据集中单个视频广告的长度6~30s,视频使用PAL制式,单个广告包含的视频帧数范围是从150~750帧.为了模拟媒体广播播出的视频流,使用视频编辑软件随机对原始的广告视频片段进行了拼接组合.在视频编辑的过程中对一半的视频使用了快进和慢动作的效果进行了处理.实验结果如表2所示.

选取不同的阈值,将影响到最终的视频比对结果,在阈值为128时,识别率较低,加大阈值可以有效提高识别率,而当阈值取2048时,可能出现的过匹配现象而出现漏检,综合考虑实验结果,可以选择512到1024为合适的阈值,此时的识别率在95%以上,能满足实际应用的要求.

表 2 改进的动态时间规划算法实验结果

Table 2 The improved DTW algorithm experiment results

阈值	正确检索	误检	漏检	查全率/%	准确度/%
128	127	67	6	95	63.5
512	189	11	0	100	94.5
2 048	199	0	1	99.5	99.5

## 4 结束语

针对电视信道中存在的 3 种主要干扰情况对广告视频检索提出改进,用视频时序特征的变化趋势取代视频时序特征值本身进行比对,并引入了异常因子用于处理突变干扰情况.本算法在将准确率控制在可容忍范围内的同时大幅度提高了查全率.本文提出的改进的动态时间规划算法能有效克服视频再编辑过程中出现的帧率变化问题.实验结果证明该算法适应性广,且执行效率得到提高.

### 参考文献:

- [1] NAPHADE M R, BASU S, SMITH J R. A statistical modeling approach to content based video retrieval[J]. Proceedings-International Conference on Pattern Recognition, 2002, 16(2): 953-956.
- [2] AL-OMARI F A, AL-JARRAH M A. Query by image and video content: a colored-based stochastic model approach[J]. Data & Knowledge Engineering, 2005, 52(3): 313-332.
- [3] PICKERING M J, RUGER S. Evaluation of key frame-based retrieval techniques for video[J]. Computer Vision and Image Understanding, 2003, 92(2~3): 217-235.
- [4] HAMPAPUR A, BOLLE R. Comparison of sequence matching techniques for video copy detection[C]//Proc. of Conf. on Storage and Retrieval for Media Databases. San Jose, USA: SPIE, 2002: 194-201.
- [5] OOSTVEEN J, KALKER T, HAITSM A J. Feature extraction and a database strategy for video fingerprinting[C]//Proc. of Int. Conf. on Visual Information and Information Systems. London: Springer-Verlag, 2002: 117-128.
- [6] AGHBARI Z, KANEKO K, MAKINOCHI A. VST-model: a uniform topological modeling of the visual-spatio-temporal video features[C]//Proceedings IEEE International Conference on Multimedia Computing and Systems. Japan: IEEE CNF, 1999, 2: 163-168.
- [7] LI Y, WEN C L, XIE Z, et al. Synchronization of batch trajectory based on multi-scale dynamic time warping[C]//International Conference on Machine Learning and Cybernetics. Xi'an: IEEE CNF, 2003, 4: 2403-2408.
- [8] LIAN N X, TAN Y P, CHAN K L. Efficient video retrieval using shot clustering and alignment[C]//Proc. of the 4th International Conference on Information, Communications & Signal and Pacific-Rim Conference on Multimedia. Singapore: IEEE CNF, 2003, 3: 1801-1805.

## Video Similarity Matching Algorithm Based on Spatiotemporal Feature

JIA Ke-bin, DENG Zhi-pin, ZHUANG Xin-yue

(College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100022, China)

**Abstract:** This paper targeted at the offset of brightness, abrupt intense disturbance and problems caused by video re-editing. Video Spatiotemporal Feature was defined to represent the content of video. The proposed Grads matching algorithm and improved Dynamic Time Warping algorithm(DTW) were introduced to video similarity matching. Experimental database included 80 advertising videos and 200 re-editing videos. Results demonstrate the effectiveness of the algorithm.

**Key words:** video signal processing; content-based retrieval; video on demand; spatiotemporal feature; gradient methods; dynamic programming

(责任编辑 张士瑛)