

基于运动补偿和自适应双阈值的镜头分割

张玉珍 杨明 王建宇 戴跃伟

(南京理工大学自动化学院 南京 210094)

摘要 镜头分割是视频检索的结构化基础。提出一种高效的镜头分割算法,首先针对视频中的图像帧序列间存在一定的运动,对视频中的图像帧采用运动矢量场自适应搜索技术进行分块运动矢量估计,然后进行运动补偿,接着在此基础上计算两帧间的像素值不变点数,并基于滑动窗口获取局部自适应的双阈值,最后基于自适应双阈值间的比较检测切变和渐变。实验结果证实,该算法不仅能够有效地检测出镜头的切变和渐变,而且对运动具有较强的鲁棒性。

关键词 分块运动矢量估计,运动补偿,自适应双阈值,镜头分割

Shot Segmentation Based on Motion Compensation and Self-adaptive Dual Thresholds

ZHANG Yu-zhen YANG Ming WANG Jian-yu DAI Yue-wei

(School of Automation, Nanjing University of Science & Technology, Nanjing 210094, China)

Abstract Shot segmentation is structural base for video retrieval. An efficient shot segmentation algorithm was proposed in this paper. Firstly aiming at the motion among the video frames, motion vector field adaptive search technique was used to estimate block motion vectors between two video frames and then motion compensation was made. Afterwards on this base the number of pixels whose value don't change much between two frames were computed and then based on sliding window, local self-adaptive dual thresholds were gained. Finally based on comparison between dual thresholds, cuts and gradual transition were detected. Experimental results prove that this algorithm can not only efficiently detect cuts and gradual transitions, but also is robust to motion.

Keywords Block motion vector estimation, Motion compensation, Self-adaptive dual thresholds, Shot segmentation

1 引言

广告是商家宣传自己产品的重要手段,夹杂在影视视频中的广告其播放时间只是影视视频中极短的一部分。为了快速地从海量的视频数据中找出商家感兴趣的广告视频,广告视频检索已成为一个研究热点。视频数据是非结构化数据,对视频数据的结构化是实现视频数据检索的前提。而镜头分割是视频结构化检索的基础。通过它可以得到连续视频流的基本组成成分即镜头。镜头是摄影机的一次操作所拍摄的内容相似的视频图像,任何一段视频数据流都是由许多镜头组成的。镜头可以分为切变(Cut)和渐变(Gradual Transition)两种。切变指发生在单一帧上的突然变化。渐变则是从一种镜头逐步过渡到另外一种镜头。镜头分割主要是根据镜头内部的相邻帧之间具有很大的相似性,当发生镜头边界时,该相似性被破坏。对于切变和渐变,这种相似性的破坏程度不同。

现有镜头检测方法大致分为两类:一是基于机器学习的方法,如文献[1,2]。在这种方法中,或者通过视频变化数学模型的推导去检测视频边界,或者通过训练得到视频变换的特有统计模型去实现视频变换检测。这种方法能有效检测出渐变,但是复杂度较大。另一种是基于阈值方法。这种方法相对比较简单,程序上容易实现。对于切变和渐变,如果使用一个阈值,难以同时检测出切变和渐变,如果阈值太小,会出

现过检测的情况,阈值太大,则难以检测出渐变的镜头边界^[3],如文献[4]基于自适应的单阈值只能有效地检测出新闻视频的切变,而渐变检测效果较差。因此为有效地检测出渐变和切变,必须使用两个阈值。Zhang等在文献[5]中最早提出镜头渐变检测的双阈值比较的方法,但是其阈值是预先设定的全局固定阈值。实际中,同一部视频的颜色分布常会随视频中的图像帧变化而变化,所以对于同一个视频不宜从头到尾使用全局固定阈值进行镜头分割。文献[6]在文献[5]的基础上提出根据 x^2 直方图获取局部自适应双阈值,实现镜头检测,但是因为 x^2 直方图的求取机制,发生帧数较长的渐变会被漏检掉。

广告商为了加强广告效应和节省时间,广告视频中编辑手法多变,镜头变化速度较快,相比其他视频,其镜头变换比较难以检测。但是无论如何,当视频发生镜头变换时,相邻图像帧中对应的像素点会发生较明显的变化;另外在视频帧之间存在一定的运动。因此本文在文献[5]的基础上,提出了一种基于运动补偿的图像像素值不变点的广告视频镜头检测方法。

像素值不变点方法主要是以前后两帧之间像素值改变不大的像素点数目作为镜头变化的度量特征,然后和阈值比较以判断镜头变化。该方法对于运动比较敏感。针对这点,可以先估计两帧间的分块运动矢量,然后进行运动补偿,再计算

到稿日期:2009-10-16 返修日期:2010-01-17 本文受南京理工大学科技发展基金(XKF09023)资助。

张玉珍(1973-),女,博士生,讲师,主要研究方向为基于语义的视频检索、图像处理及模式识别,E-mail:olindazh@163.com。

经运动补偿过的两帧间的像素值不变点数,最后基于自适应双阈值实现镜头分割。

2 分块运动矢量估计

分块运动估计方法是把视频图像帧按照 $M \times N$ 的大小分成一定数目的子块,然后利用块匹配算法匹配子块以得到平移矢量。块匹配算法(Block Matching Algorithm, BMA)是根据子块平移运动的机理进行运动估计,在各类的运动估计算法中,由于其简单高效、额外开销较小因而被广泛使用。分块运动估计原理如下:先从视频中取出两帧相邻的图像,设前一帧为参考帧,后一帧为当前帧,把当前帧分成多个二维的图像子块,假定子块内每个像素都有同样的运动矢量,并且只进行平移运动,不进行旋转运动,取出其中一个子块为当前子块,并在参考帧同样位置四周一定范围(该范围也称为搜索窗口)内按照匹配准则寻找与其最接近的子块,称与其最接近的子块为预测块,当前块和预测块之间的位移称为运动矢量,依次估计当前帧所有子块的运动矢量就是分块运动矢量估计。如图 1 所示, (dX, dY) 即为当前子块的运动矢量,运动矢量是通过固定的当前块搜索可变的预测块,因此图中运动矢量的方向是从当前块指向预测块,这在时间顺序上是逆向的。

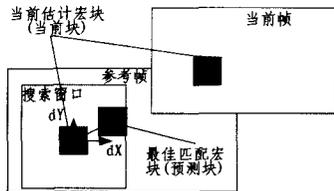


图 1 分块运动矢量估计示意图

分块运动矢量估计需要解决两个问题,即匹配准则和运动估计方式(也称为搜索定位方法)。其中,匹配准则中为了计算方便,通常可采用绝对误差和 SAD(Sum of Absolute Difference),公式如下:

$$SAD(i, j) = \sum_{m=1}^M \sum_{n=1}^N |G_x(m, n) - G_{x-1}(m+i, n+j)| \quad (1)$$

式中,子块宽度和高度分别为 M 和 N 。 (i, j) 为位移量, (m, n) 为像素点在子块中的坐标。 $G_x(m, n)$ 是一个当前帧像素点的像素值,像素点在当前帧中的坐标为当前子块相对于当前帧的偏移量加上 (m, n) , $G_{x-1}(m+i, n+j)$ 是同一个点在参考帧中的像素值,该点在参考帧中的坐标为当前子块相对于当前帧的偏移量加上 $(m+i, n+j)$ 。

根据搜索定位方法搜索 (i, j) ,若在 (i_0, j_0) 处 SAD 达到最小,则 (i_0, j_0) 即为计算得到的位移估计量。

本文的搜索定位方法采用文献[7]中的运动矢量场自适应搜索技术,且以文献[8]中的菱形搜索法为基础。

2.1 菱形搜索算法

菱形搜索法^[8](Diamond Search, DS)采用了大菱形模版(Large Diamond Search Pattern, LDSP)和小菱形模版(Small Diamond Search Pattern, SDSP)进行搜索匹配。大菱形模版是搜索如图 2(a)所示的 9 个点(每个点就是一个位移量 (i, j)),根据匹配准则找出最优矢量的搜索点,如果这个搜索点在模版中心,则转入小菱形模版搜索,否则以这个搜索点为中心继续采用大菱形模版搜索。小菱形模版是搜索如图 2(b)所示的 5 个点的位置,根据匹配准则找出最优矢量的搜索点,如果这个搜索点在模版中心,则认为找到最优矢量的搜索点,

否则以这个搜索点继续进行小菱形模版搜索。图 3 给出了一个搜索过程实例,其中搜索窗口大小为 15×15 ,起始搜索中心为 $(0, 0)$ 。该例以 5 个搜索步骤给出最终的最优矢量搜索点 $(-4, -2)$;4 次大菱形模版搜索和最后的 1 次小菱形模版搜索。在该例中总共有 24 个搜索点:即依次在每个步骤中分别搜索 9 个点、5 个点、3 个点、3 个点及 4 个点。

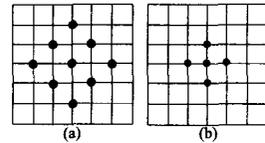


图 2 大菱形模版和小菱形模版

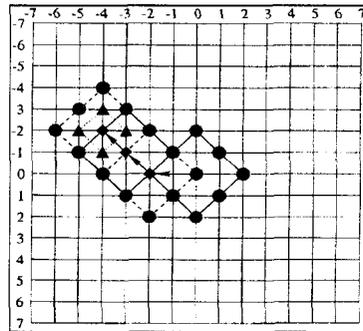


图 3 搜索过程实例

2.2 运动矢量场自适应搜索技术

运动矢量场自适应搜索技术^[7](Motion Vector Field Adaptive Search Technique, MVFAST)以菱形搜索法为基础,加入了预测算法,根据上一帧的运动矢量估计结果预测运动类型,把运动类型分为高、中、低 3 种类型。区分方法如下:根据上一帧中当前块和其左面、上面、右上面 3 块的运动矢量来估计当前帧中当前块的运动矢量,如图 4 所示。设这 4 块运动矢量分别为 V_0, V_1, V_2, V_3 。其中 $V_i = (x_i, y_i)$, $i=0, 1, 2, 3, V_0 = (0, 0)$, 设 $L_i = |x_i| + |y_i|$, $L = \text{MAX}(L_i)$ 。设有阈值 $L1 \leq L2$, 若 $L \leq L1$, 则称当前块运动类型为低(速);若 $L1 < L \leq L2$, 则称当前块运动类型为中(速);若 $L2 < L$, 则称当前块运动类型为高(速)。



图 4 用来矢量估计的相邻块示意图

当运动类型为中时,采用的搜索方式为先采用大菱形搜索,再进行小菱形搜索(这种搜索方式简称为 LDS),其余运动类型则直接进行小菱形搜索(这种搜索方式简称为 SDS)。当如果运动类型为高时,搜索窗口的搜索中心(search center)为计算图 4 中几个子块的运动矢量产生 SAD 最小的点,其余类型的搜索中心直接为 $(0, 0)$ (即当前块不移动)。其中大菱形搜索(LDS)和小菱形搜索(SDS)步骤分别如下所述。

i) 小菱形搜索(SDS)

Step 1 以搜索中心为中心定位小菱形模版,依次搜索小菱形模版的所有点,根据匹配准则 SAD 找出最优矢量的搜索点,如果这个搜索点在模版中心,则这个点就是要找的最优分块运动矢量,否则,转到 Step 2。

Step 2 小菱形模板的中心移到在上个步骤中找到的搜索点(即 SAD 最小对应的点),依次搜索小菱形模板的所有点,根据匹配准则 SAD 找出最优矢量的搜索点,如果这个搜索点在模板中心,则这个点就是要找的最优分块运动矢量,否则,就重复 Step 2。

ii) 大菱形搜索(LDS)

Step 1 以搜索中心为中心定位大菱形模板,依次搜索大菱形模板的所有点,根据匹配准则 SAD 找出最优矢量的搜索点,如果这个搜索点在模板中心,就转到 Step 3,否则转到 Step 2。

Step 2 大菱形模板的中心移到在上个步骤中找到的搜索点(即 SAD 最小对应的点),依次搜索大菱形模板的所有点,根据匹配准则 SAD 找出最优矢量的搜索点,如果这个搜索点在模板中心,就转到 Step 3,否则重复 Step 2。

Step 3 搜索模板由大菱形模板转到小菱形模板。依次搜索小菱形模板的所有点,产生 SAD 最小的点就是要找的最优分块运动矢量对应的点。

3 广告镜头分割算法

3.1 经过运动补偿的像素值不变点的计算

像素值不变点是指前后两帧之间像素值改变不大的像素点数目。因为在视频的图像序列中,存在背景和物体的运动,所以为提高检测效率需要采用一定的运动补偿。本文采用第 2 节所述的 MVFAST 分块运动矢量估计方法,通过当前帧和参考帧,计算出分块运动矢量,接着利用参考帧和运动矢量根据式(2)重构出预测帧,再对预测帧和当前帧根据式(3)进行比较得到相邻帧间经过运动补偿后的像素值不变点数。

$$P'_i(x, y) = P_i(x + dx, y + dy) \quad (2)$$

式中, $P_i(x + dx, y + dy)$ 为第 i 帧图像在坐标 $(x + dx, y + dy)$ 处的像素值, (x, y) 为第 $i + 1$ 帧中一像素点的坐标, (dx, dy) 为第 $i + 1$ 帧与第 i 帧之间该点所对应的分块运动矢量, $P'_i(x, y)$ 为预测帧在坐标 (x, y) 处的像素值。

$$Q_{i+1,i} = \sum_{x=1}^X \sum_{y=1}^Y f_{i+1,i}(x, y)$$

$$f_{i+1,i}(x, y) = \begin{cases} 0, & g_{i+1,i}(x, y) > t \\ 1, & g_{i+1,i}(x, y) < t \end{cases}$$

$$g_{i+1,i}(x, y) = |P_{i+1}(x, y) - P'_i(x, y)| \quad (3)$$

式中, X, Y 是图像的长度和宽度, $P_{i+1}(x, y)$ 是指第 $i + 1$ 帧(设为当前帧)中坐标为 (x, y) 像素点的像素值, $P'_i(x, y)$ 为根据第 i 帧(即参考帧)及第 $i + 1$ 帧与第 i 帧之间的分块运动矢量重构出的预测帧在坐标为 (x, y) 像素点的像素值。 $Q_{i+1,i}$ 表示第 i 帧与经过运动补偿后的预测帧间像素值不变点数。 t 为判断不变点的阈值。 t 主要用来抑制噪声,如信号转换或者编码产生的随机噪声,一般取 3~5 即可。

加运动补偿的像素值不变点特征能够更好地反映出两幅图像的差异,对原有特征造成的影响很小。如图 5 所示,参考帧和当前帧是一个镜头内的相邻帧,本来差异不大(特征值 $Q_{i+1,i}$ 为 71350),经过运动补偿后差异仍然不大(特征值 $Q_{i+1,i}$ 为 90991)。如图 6 所示,参考帧和当前帧是镜头切变处的相邻帧,本来差异很大(特征值 $Q_{i+1,i}$ 为 35591),经过运动补偿后差异仍然很大(特征值 $Q_{i+1,i}$ 为 38241)。而图 7 是一个镜头内的相邻帧,人眼可以看出两幅图差异不是很大,但是由于运动块较多,直接应用像素值不变点特征显示差异很大

(特征值 $Q_{i+1,i}$ 为 28883),经过运动补偿后差异小了很多(特征值 $Q_{i+1,i}$ 为 60071),运动补偿取得了较好的效果。



图 5 相邻帧序列运动补偿图例 1



图 6 相邻帧序列运动补偿图例 2



图 7 相邻帧序列运动补偿图例 3

当发生镜头边界时,相邻帧中对应的像素点会发生明显变化,因为边界转换机制不同,其中渐变时的像素点变化要明显小于切变时的,即相邻帧间的像素值不变点数明显大于切变时的,而且这种变化是一个逐步进行的过程,会持续若干帧。因此在基于运动补偿过的像素值不变点特征检测镜头边界时,需要设置两个阈值 T_l 和 T_h 来区分切变和渐变,其中 $T_l < T_h$ 。当相邻帧间的经过运动补偿后的像素值不变点数 $Q_{i+1,i} \leq$ 低阈值 T_l 时,认为发生切变,而当 $T_l < Q_{i+1,i} \leq T_h$ 时,认为该两帧中的后一帧为可能存在的渐变的起始帧。然后同时计算两种经过运动补偿过的像素值不变点数:一种是相邻两帧间的经过运动补偿后的像素值不变点数 $Q_{i+1,i}$;另一种是称为相隔帧间的像素值不变点数 Q_n ,即经过运动补偿后的后续第 n 帧与潜在起始帧间的像素值不变点数。因为在渐变结束时图像帧(即渐变尾帧)与起始帧间的像素值不变点数将具有突变的特征,所以渐变结束时所对应的经过运动补偿后的相隔帧间的像素点不变数(即渐变尾帧与渐变起始帧间的像素值不变点数)应该不超过较低的阈值 T_l 。因此当经过运动补偿后的相隔帧间的像素值不变点数低于低阈值 T_l 而经过运动补偿后的相邻帧间的像素值不变点数高于高阈值 T_h 时,对应帧便为镜头渐变的终止帧。另外,因为在整个渐变过程中,起始帧将不断被另一个镜头的画面所代替,所以起始帧的特征将越来越小,而渐变结束时的图像帧的特征将越来越大,因此在整个渐变过程中, Q_n 是单调递减的。整个过程如图 8 所示。

计算经过运动补偿后的相隔帧间的像素值不变点数的公式为

$$Q_{q+n,q} = \sum_{x=1}^X \sum_{y=1}^Y f_{q+n,q}(x, y)$$

$$f_{q+n,q}(x, y) = \begin{cases} 0, & g_{q+n,q}(x, y) > t \\ 1, & g_{q+n,q}(x, y) < t \end{cases}$$

$$g_{q+n,q}(x, y) = |P'_{q+n-1}(x, y) - P_q(x, y)| \quad (4)$$

式中, $P_q(x, y)$ 是指第 q 帧中坐标 (x, y) 为像素点的像素值, $P'_{q+n-1}(x, y)$ 为根据第 $q + n - 1$ 帧及第 $q + n$ 帧与第 $q + n - 1$ 帧之间的分块运动矢量重构的预测帧在坐标为 (x, y) 像素点

的像素值。

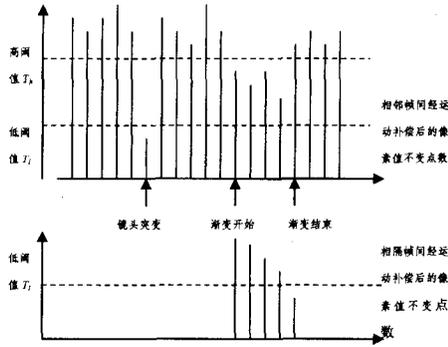


图8 基于双阈值比较的镜头边界检测示意图

3.2 自适应阈值 T_l 和 T_h 的计算

由于同一视频中各镜头边界的变化是不固定的,因此对于同一个视频不宜从头到尾使用全局固定阈值进行镜头分割,而必须随视频中图像帧的变化采用自适应局部阈值检测镜头边界。

自适应局部阈值的获取思想为:设计一个滑动窗口,滑动窗口的宽度 W 初始化为 1 帧,增长步长为 1 帧。根据式(3)逐个计算滑动窗口内相邻两帧间的像素值不变点数 $Q_{i+1,i}$ 。然后根据式(5)和式(6)计算滑动窗口内的相邻帧间的像素值不变点数的平均值 u 和标准方差 σ 。

$$u = \frac{1}{W} \sum_{i \in W} Q_{i+1,i} \quad (5)$$

$$\sigma^2 = \frac{1}{W} \sum_{i \in W} (Q_{i+1,i} - u)^2 \quad (6)$$

根据均值 u 和标准方差 σ 及式(7)和式(8)计算两个自适应阈值 T_l 和 T_h , 其中 $T_l < T_h$ 。

$$T_h = \lambda_h (\mu - n\sigma) \quad n=1 \sim 3, \lambda_h = 0.8 \sim 1.0 \quad (7)$$

$$T_l = \lambda_l \mu \quad \lambda_l = 0.3 \sim 0.5 \quad (8)$$

3.3 广告镜头分割算法

广告镜头分割算法步骤如下:

(1)初始化滑动窗口 W 的宽度为 1 帧。

(2)以步长为 1 增长滑动窗口,并根据第 2 节所述的 MVFAST 分块运动矢量估计方法及式(2)、式(3)、式(5)一式(8)计算滑动窗口内的经过运动补偿过的相邻两帧间像素值不变点数、均值 u 、标准方差 σ 及阈值 T_l 和 T_h ,若滑动窗口内末两帧间的经过运动补偿后的像素值不变点数 Q_{last} 大于高阈值 T_h ,则重复本步骤,否则转到步骤(3)。

(3)若经过步骤(2)计算得到的 $Q_{last} \leq T_l$,则认为发生镜头切变,即此时滑动窗口的倒数第二帧是上一个镜头的尾帧,而倒数第一帧则为下一个镜头的起始帧,标记该切变。然后重新初始化滑动窗口宽度 W 为 1,窗口内包含该起始帧,并转到步骤(2)继续寻找下一个镜头边界;否则转步骤(4)。

(4)若经过步骤(2)计算得到的 Q_{last} 在 T_l 和 T_h 之间,即 $T_l < Q_{last} \leq T_h$,则认为此时滑动窗口的最后一帧为可能存在的镜头渐变的起始帧,记这一帧为 q ,以步长 1 循环增长滑动窗口并根据第 2 节所述的 MVFAST 分块运动矢量估计方法及式(2)一式(8)分别计算经运动补偿后的相邻两帧间的像素值不变点数 $Q_{k+1,k} (k=q, q+1, q+2, \dots, q+m-1, m$ 是该渐变过程的最大帧数)、经运动补偿后的相隔帧间的像素值不变点数 $Q_n (Q_n = Q_{q+n}, n=1, 2, 3, \dots, m)$ 、均值 u 、标准方差 σ 及阈值 T_l 和 T_h ,当经运动补偿过的相隔帧间的像素值不变点数

$\leq T_l$ 而经运动补偿过的相邻帧间的像素值不变点数 $> T_h$,则此时滑动窗口的最后一帧便为镜头渐变的尾帧,标记该渐变,然后重新初始化滑动窗口宽度 W 为 1,窗口内包含该尾帧的下一帧,并转到步骤(2)继续寻找下一个镜头边界。注意在 Q_n 开始计算后,同时观察 $Q_{n+1,n}$ 和 Q_n ,如果 Q_n 不是单调递减,或如果经运动补偿过的相邻帧间的像素值不变点数 $> T_h$,而此时经运动补偿过的相隔帧间的像素值不变点数 $> T_l$,则废除该渐变,转到步骤(2)继续寻找镜头边界。

4 实验结果及分析

为检测本文提出的镜头分割算法的有效性,进行了以下实验。实验中与文献[5,6]中的镜头检测算法进行了比较。实验环境是 VC++6.0。实验中对 3 个视频片段进行镜头分割。检测结果见表 1 和表 2。其中广告片段 1 和广告片段 2 都是从 CCTV1 频道录制的若干个广告视频的组合。因为广告视频经常夹杂在影视视频中播放,所以为了检测本算法的有效性,实验中还包含了一段电视剧片段和一段电影片段。镜头分割的界面如图 8 中(a),(b)图所示,界面图中左视图为视频播放窗口,右视图为检测出的镜头边界处的图像帧。根据查全率和查准率对算法进行检测,其中查全率(Recall)和查准率(Precision)定义分别为:

$$\text{Recall} = \text{Correct} / (\text{Correct} + \text{Miss})$$

$$\text{Precision} = \text{Correct} / (\text{Correct} + \text{False})$$

实验数据表明本文提出的算法对于镜头切变检测效果很好,对于渐变检测也较满意。数据表明本文提出的算法比文献[5,6]的效果都要好,尤其比文献[5]高出很多。这是因为文献[5]中不仅没有考虑运动补偿,而且采用固定的全局双阈值,当其检测较长的视频片段时,效果较差。文献[6]中因为没有考虑运动补偿,以及 χ^2 直方图的求取机制,所以突变和切变的检测效果也都没有本文算法的效果好。图 9 是基于本文算法的镜头分割示意图。

表 1 镜头分割结果(查全率和查准率的单位均为%)

| 视频片段 | 镜头个数 | 文献[5]算法 | | 文献[6]算法 | | 本文算法 | |
|---------|------|---------|------|---------|-------|-------|-------|
| | | 查全率 | 查准率 | 查全率 | 查准率 | 查全率 | 查准率 |
| 广告切变 | 60 | 68.3 | 66.1 | 85.0 | 89.5 | 95.0 | 96.6 |
| 片段 1 渐变 | 16 | 62.5 | 58.8 | 75.0 | 80.0 | 87.5 | 93.3 |
| 广告切变 | 51 | 66.7 | 65.4 | 80.4 | 83.7 | 94.1 | 96.0 |
| 片段 2 渐变 | 20 | 60.0 | 57.1 | 80.0 | 84.2 | 90.0 | 94.7 |
| 电视剧切变 | 113 | 61.9 | 60.9 | 86.7 | 89.9 | 93.8 | 96.4 |
| 片段 渐变 | 10 | 70.0 | 63.6 | 80.0 | 88.9 | 90.0 | 90.0 |
| 电影切变 | 61 | 63.9 | 61.9 | 86.9 | 89.8 | 95.1 | 96.7 |
| 片段 渐变 | 5 | 60.0 | 60.0 | 80.0 | 100.0 | 100.0 | 100.0 |

表 2 镜头分割平均结果

| | 文献[5]算法 | | 文献[6]算法 | | 本文算法 | |
|----|---------|-------|---------|-------|-------|-------|
| | 平均查全率 | 平均查准率 | 平均查全率 | 平均查准率 | 平均查全率 | 平均查准率 |
| 切变 | 65.2% | 63.6% | 84.8% | 88.2% | 94.5% | 96.4% |
| 渐变 | 63.1% | 59.9% | 78.8% | 88.3% | 91.9% | 94.5% |

实验数据表明本文提出的镜头分割算法不仅适合于广告视频,而且也适合于影视视频。因为一般视频在发生镜头变换时,相邻图像帧中对应的像素点会发生较明显的变化,而且视频帧间存在运动,而本算法是在运动补偿的基础上计算两帧间的像素值不变点数,并基于滑动窗口计算局部自适应双阈值进行镜头分割,所以本文提出的镜头分割算法可以推广

到其他类视频。因为采用了运动补偿,所以对视频帧间的运动具有较强的鲁棒性,例如在广告视频片段2中有一段汽车广告的视频帧间的运动幅度较大,但是并不影响检测效果。

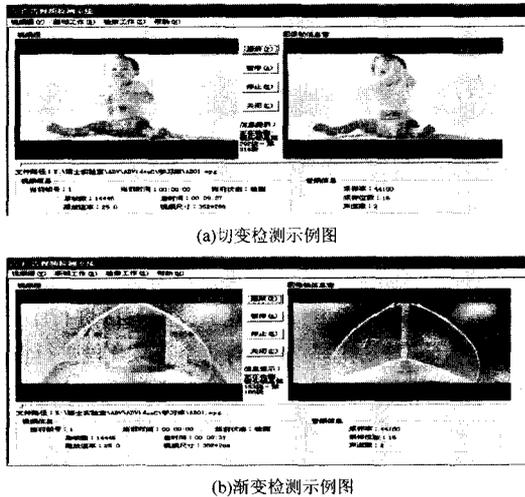


图9 基于本文算法的镜头分割示意图,其中图(a)检测出的是一个切变,图(b)检测出的是一个渐变

结束语 镜头分割是视频检索的重要的结构化基础。本文针对广告视频提出了一种镜头分割算法。首先对视频中的图像帧进行运动补偿,然后在此基础上计算两帧图像间的像素值不变点数,并基于滑动窗口获得局部自适应的双阈值,对镜头的切变和渐变进行检测。实验证实该算法不仅能有效地检测切变和渐变,而且还可以推广到其他类视频,为后续的广

告视频检索提供视频结构化基础。

参考文献

- [1] Cao J R, Cai A N. Algorithm for Shot Boundary Detection Based on Support Vector Machine in Compressed Domain[J]. Acta Electronica Sinica, 2008, 36(1): 203-208
- [2] Zhang N, Xiao G Q, Jiang J M, et al. BPNN Algorithm towards Shot Boundary Detection[J]. Journal of Computer Application, 2009, 29(5): 1369-1372
- [3] Zhi M, Cai A N. Shot Boundary Detection with Main Color[J]. Acta Automatica Sinica, 2007, 33(6): 654-657
- [4] Zhao N, Lv N, Liu H Y. Content-Based Cut Shot Detecting Algorithm of NewsVideo[J]. Journal of Jilin University: Information Science Edition, 2009, 27(1): 50-55
- [5] Zhang H J, Kankanhalli A, Smoliar S W. Automatic partitioning of full motion video[J]. Multimedia System, 1993, 1(1): 10-28
- [6] Wang J Y, Luo W. A Self-adapting Dual-threshold Method for Video Shot Transition Detection[C]//IEEE International Conference on Networking, Sensing and Control. 2008, 4: 704-707
- [7] Ma K K, Hosur P I. Performance Report of Motion Vector Field Adaptive Search Technique (MVFAST) [R]. MPEG2000/M5851. 2000
- [8] Zhu S, Ma K K. A new diamond search algorithm for fast block matching motion estimation[J]. IEEE Transactions on Image Processing, 2000, 9(2): 287-290
- [9] Zhuang Y T, Pan Y H, Wu F. Web-based Multimedia Information Analysis and Retrieval[M]. Beijing: Tsinghua University Publishing House, 2002, 9: 76
- [10] 熊江. 基于小波变换的图像去噪[J]. 计算机科学, 2007, 34(7): 232-234
- [11] Pajares G. A wavelet based image fusion tutorial [J]. Pattern Recognition, 2004, 37(9): 1855-1872
- [12] 江泽民, 徐德, 杨月全, 等. 基于小波包分解的图像清晰度判定[J]. 中国图象图形学报, 2004, 9(6): 717-722
- [13] Tikkanen P E. Nonlinear wavelet and wavelet packet denoising of electrocardiogram signal [J]. Biological cybernetics, 1999, 80(4): 259-267
- [14] Jorge S, Narayanan, Shrikanth S. Discriminative Wavelet Packet Filter Bank Selection for Pattern Recognition [J]. IEEE Transactions on Signal Processing, 2009, 57(5): 1796-1810
- [15] Brechet L, Lucas M F, Doncarli C, et al. Compression of Biomedical Signals With Mother Wavelet Optimization and Best-Basis Wavelet Packet Selection [J]. IEEE Transactions on Biomedical Engineering, 2007, 54(12): 2186-2192
- [16] Cristan A C, Walden A T. Multitaper Power Spectrum Estimation and Thresholding: Wavelet Packets Versus Wavelets [J]. IEEE Transactions on Signal Processing, 2002, 50(12): 2976-2986
- [17] Ercelebici E, Koc S. Lifting-based wavelet domain adaptive Wiener filter for image enhancement [J]. IEE proceedings. Vision, image and signal processing, 2006, 153(1): 31-36
- [18] Hu Shaoxiang, Liao Zhiwu. Signal denoising using wavelet packet hidden Markov model [C]//2004 International Conference on Communications, Circuits and Systems. Wuhan, 2004: 751-755
- [19] Toliyat H A, Abbaszadeh K, Rahimian M M, et al. Rail Defect Diagnosis Using Wavelet Packet Decomposition [J]. IEEE Transactions on Industry Applications, 2003, 39(5): 1454-1461

(上接第 281 页)

图像中都存在一些明显的噪点未能消除。图 8 中右边焊点,去噪效果相对要好一些,但对于左边焊点,去噪效果没有采用本文提出的方法理想。

结束语 提出了一种基于小波包变换结合 wiener 滤波的 SMT 焊点图像去噪新方法。采用小波包分解算法对 SMT 焊点图像逐层进行分解,其不仅对低频段进行分解,而且对高频段也进行分解。对这些小波包系数进行相应处理,有利于更好地实现 SMT 焊点图像的去噪。

结合 Wiener 滤波,对小波包系数进行合理处理,比直接采用小波包系数阈值处理方法更合理。与传统的方法相比,去噪结果更为理想,同时能较好地保存图像的边缘信息。由于采用了中值绝对方差估计方法对原图像的噪声进行估计,因此本文提出的方法对存在高斯噪声的 SMT 焊点图像去噪具有一定的通用性。同时,由小波包变换和 Wiener 滤波的去噪的特点决定了,该方法在去除其它类型的噪声上也有一定的效果。

由于 SMT 焊点类型多,并且各种原因产生的噪声类型也有较大的差别,要将本文提出的方法上形成实用系统,还有很多工作要做。

参考文献

- [1] 周德俭,黄春跃,吴兆华,等. SMT 焊点虚拟成形和 SMT 产品虚拟组装技术研究[J]. 计算机集成制造系统, 2006, 12(8): 1267-1272
- [2] Kong Fanhui, Wang Yongxin. Reconstruction of solder joint surface based on shape from shading [C]//IEEE Third International Conference on Natural Computation. New York, 2007: 58-