

一种基于灰度连续区域分割的 视频对象分割方法*

陈韩锋 戚飞虎

(上海交通大学, 计算机科学与工程系, 上海, 200030)

摘要 针对目前许多分割方法中分割边界不精确、计算复杂和缓存帧多等问题, 提出了一种结合空间区域分割和运动像素检测的自动分割方法: 先将当前视频帧分割为不同的灰度连续区域, 再利用二次帧差确定视频图像中的运动像素, 然后按一定的规则确定哪些灰度连续区域属于运动区域, 从而有效地从静止的复杂背景中分割出运动对象区域. 实验结果表明这种分割方法计算简单、分割边界比较精确.

关键词 MPEG-4, 视频对象分割, 灰度连续区域, 帧差图.

VIDEO OBJECT SEGMENTATION BASED ON GREY-CONSECUTIVE REGIONS SEGMENTATION

CHEN Han-Feng QI Fei-Hu

(Department of Computer Science & Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract A spatio-temporal method which aims at segmenting moving objects from stationary background automatically and accurately was proposed. Three steps are followed steps in the algorithm. First, the current frame is segmented into connected regions of uninterrupted grey scale. Secondly, The intersections of each two neighbor difference images are gotten to judge which pixels are moving-pixels. Finally, those moving regions are segmented from the stationary regions according to a certain rule. The experimental results show that the algorithm is effective in the segmenting moving objects from the stationary background accurately.

Key words MPEG-4, video object segmentation, grey-consecutive region, difference image.

引言

视频对象分割技术是 MPEG-4 标准中的一项关键技术^[1]. 视频对象分割是指将输入视频的每一帧分割为不同语义(semantic)的对象区域. 本文的研究重点是从静止的复杂背景中分割出视频中运动的前景对象区域(前景运动对象区域).

目前的视频对象分割技术主要有半自动分割技术^[2,3]和自动分割技术^[4-6]. 半自动分割技术在分割过程中, 借助人机交互提取人感兴趣的视频对象, 这种方法的主要缺点是不适合对分割有实时性要求的场合, 而且操作人员的负担太重. 自动分割技术主要是利用帧与帧之间由于运动引起的像素变化和帧内的空间聚类信息(如灰度、纹理等), 分割出感兴趣的对象区域. 自动分割可利用的信息相对较少, 易受

噪声的影响, 分割的边界比较粗糙. 为了增强抗噪能力, 有人利用多帧的统计信息^[4]或高阶统计检测方法^[5,10]来确定像素的运动信息, 这种方法在分割当前帧的时候, 需要缓存许多相邻帧, 而且计算比较复杂. 文献[6]提出一种运动区域的分割和跟踪方法, 先用分级形态分割算法^[7]将每帧分割为许多区域, 然后利用两个相邻帧估计各区域的运动矢量, 最后合并各个区域得到运动区域, 由于运动模型估计的误差, 这种方法的分割边界不很令人满意.

本文提出一种结合利用视频帧的空间域分割信息和时间域分割信息的视频对象自动分割方法, 从静止的复杂背景中分割出视频中运动的前景对象区域. 这种方法先根据灰度信息将每帧视频图像分割成不同的连通区域, 然后根据相邻帧之间的运动信息, 确定每个连通区域是否属于运动视频对象区域.

* 国家自然科学基金(编号 60072029)资助项目

稿件收到日期 2001-06-06, 修改稿收到日期 2001-07-31

* The project supported by the National Natural Science Foundation of China (No. 60072029)

Received 2001-06-06, revised 2001-07-31

1 基于灰度连续区域分割的运动目标分割算法

一个视频对象对应视频图像中一定的对象区域,而每个对象区域总是由一些连通区域构成的.这些连通区域可以是单灰度区域^[8](同一个区域内的所有像素具有相同灰度值),也可以是灰度连续区域(区域内的相邻像素具有相似的灰度值).若找到这些连通区域,分割出这些对象区域,也就分割出了视频对象,但在一帧复杂图像中,仅仅利用空间域信息自动分割出具有语义含义的对象区域是很困难的.

利用视频序列中像素点在帧与帧之间的运动信息,可以将运动的前景对象 (foreground) 从背景 (background) 中分割出来. 像素的运动信息可以通过帧差图像表现出来,或通过块匹配方法得到其运动矢量. 然而,仅仅利用像素在时间域上的变化信息来分割运动视频对象也存在许多问题:首先是对噪声的敏感性,图像噪声的存在会使一些背景区域被误认为是运动的前景对象区域;其次,由于前景对象与背景之间的遮挡和显露关系,当前景对象运动后显露出的背景区域会被误认为是运动的前景区域;更严重的是,如果前景区域中存在大面积具有相同颜色的区域,则该区域中大量像素的运动信息不能正确地获得,更不能在帧差图像上表现出来. 如图 1,实线多边形内为一个同灰度区域,在下一帧中,该对象运动到了虚线框内的区域,两个区域的重叠部分在帧差图像中会留下一个大空洞(阴影部分),被误认为是静止的背景区域.

针对这些问题,本文提出了一种基于灰度连续区域的分割方法,其原理如图 2 所示. 这种方法主要由两部分组成:即灰度连续区域分割和运动区域检测. 利用帧间的像素运动信息,可以检测运动区域;利用帧内的灰度连续信息,可提高分割的精确性.

1.1 预处理

预处理主要是对视频图像进行滤波处理. 视频

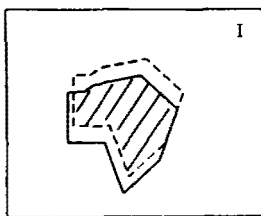


图 1 错误误的背景区域
Fig. 1 Error background

图像中不可避免的会有一些随机噪声,影响灰度连续区域的分割和运动区域的检测. 本文中采用了中值滤波的方法,在去除噪声的同时能较好地保留对象区域的边缘信息.

1.2 灰度连续区域分割

灰度连续区域的定义为:设 R 为图象 I 中的一个连通区域, R 中任意一个像素 p 都至少存在一个相邻像素 q ,使得 $|grey(p)-grey(q)| < 2$,函数 $grey(p)$ 表示像素 p 的灰度值;同时,对于 R 中的任意像素 p ,其满足 $|grey(p)-grey(q)| < 2$ 的相邻像素 q 都在 R 内,则称 R 为一个灰度连续区域. 本文中的相邻像素采用上下左右四相邻像素.

对象区域边界两侧的像素分别属于背景区域和前景区域,通常情况下,可以假定背景区域和前景区域之间像素的灰度值有突变,否则视觉无法辨别背景区域和前景区域. 因此,可以认为对象区域的边界必定是灰度连续区域的边界,即对象区域是由一些灰度连续区域组成的. 当然,灰度连续区域的边界不一定是对象区域的边界,如图 3 所示(实线为对象区域边界,虚线为灰度连续区域边界). 将视频图像分割成不同的灰度连续区域,有助于前景对象区域的分割:只要确定哪些灰度连续区域是属于前景对象区域的,就可以得到所要的对象区域,这样得到的对象区域具有“像素级精度”^[4].

本文采用了一种区域增长的方法分割灰度连续区域:即首先找到区域增长的种子(起始点),以该种子为起点判断其相邻像素是否属于同一个灰度连续区域,然后递推搜索图像中所有属于该灰度连续区域的像素点;当搜索完一个灰度连续区域后,重新寻找新的种子,搜索新的区域. 种子可以是不属于任何已发现的灰度连续区域的像素点. 递推过程可以描

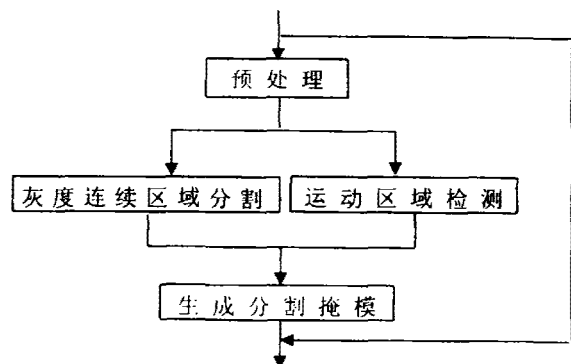


图 2 基于灰度连续区域分割的视频对象分割方法
Fig. 2 Diagram of video object segmentation based on grey-consecutive regions segmentation

述为:

neighbour(i, j)

if(像素 $p(i, j)$ 的相邻像素 $q(m, n)$ 还未加入到任何灰度连续区域且 $(|\text{grey}(p) - \text{grey}(q)| < 2)$)

将像素 $q(m, n)$ 加到灰度连续区域 R 中;
neighbour(m, n); }

1.3 运动区域检测

视频图像中的运动区域是运动对象在视频窗口内的投影,对象的运动引起运动区域内像素点的灰度值变化,这种变化可以在两帧相邻视频图像的帧差图上反映出来.帧差图经过二值化处理,就得到一个尺寸与帧差图相同的用于标志运动区域和静止区域的分割掩模,分割掩模中的 0 表示对应位置的像素点属于静止区域,1 表示对应位置的像素点属于运动区域.早期的一些运动区域分割方法就是基于这种二值化的帧差图^[9],二值化的帧差图可以描述为

$$d_k(i, j) = \begin{cases} 1, & |\text{grey}_k(i, j) - \text{grey}_{k-L}(i, j)| > T \\ 0, & |\text{grey}_k(i, j) - \text{grey}_{k-L}(i, j)| \leq T \end{cases} \quad (1)$$

其中 $d_k(i, j)$ 为二值化的帧差图, (i, j) 为像素点位置, $\text{grey}_k(i, j)$ 为第 k 帧中像素 (i, j) 的灰度值, T 是二值化所需的阈值,是个经验值,本文中 $T = 1$. 当对象运动较慢时,对应的运动区域在相邻帧之间的变化很小.可以用增加帧差图像的间隔的方法加

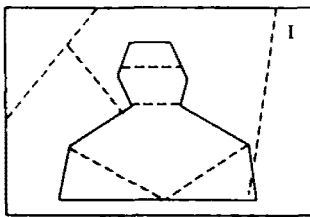


图 3 对象区域和灰度连续区域

Fig. 3 Object region and grey-consecutive region

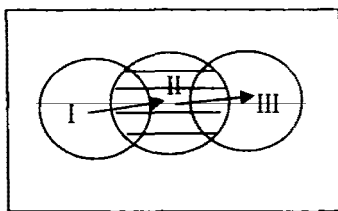


图 4 二次帧差

Fig. 4 Two-order difference image

以改善,即在检测运动区域时,求第 k 帧和 $k-L$ 帧的帧差, L 为间隔帧数.

正如第 2 节开始时所述,这种简单的差值算法存在许多问题.本文提出二次帧差方法来改善.设 d_k 和 d_{k-L} 为由式(1)得到的两个帧差图像,二次帧差 D_k 定义为

$$D_k(i, j) = \begin{cases} 1, & d_k(i, j) = d_{k-L}(i, j) = 1 \\ 0, & \text{其他} \end{cases} \quad (2)$$

这种方法用相邻的两个帧差图像的交集检测运动区域.如图 4,第 $k-L$ 、 k 帧的帧差图像检测到的运动区域为区域 I、II 的并集;第 k 、 $k+L$ 帧的帧差图像检测到的运动区域为区域 II、III 的并集;将这两个并集相交,得到区域 II,即为第 k 帧中的运动区域(阴影部分).

用这种方法检测运动区域,可以较好地避免由于遮挡与显露关系而将背景区域误认为运动的前景区域,同时有助于抑制随机噪声的影响.

1.4 生成分割掩模

分割掩模是一个宽高尺寸与视频图像相同的矩阵,矩阵的每个元素用 0 和 1 分别标志当前帧中对应位置的像素点是属于背景区域还是前景运动对象区域.生成分割掩模就是要确定哪些灰度连续区域是属于前景对象区域的.本文提出一种利用运动区域检测结果确定哪些灰度连续区域是属于前景运动对象区域的方法:检查每一个灰度连续区域的边缘像素,如果某个灰度连续区域的所有边缘像素中有超过一定比例 Ratio 的像素被检测为属于前景运动对象区域,则认为该灰度连续区域属于运动区域.具体的算法描述为

```
for( $m = 1; m \leq \text{REGION}; m = m + 1$ )
```

```
//REGION 为当前帧中灰度连续区域的个数
```

```
{for( $n = 1; n \leq \text{EDGE}(m); n = n + 1$ )
```

```
//EDGE( $m$ )为灰度连续区域  $m$  的边界点个数
```

```
{ if( $C(\text{pixel}[m][n]) = 1$ )
```

```
// pixel[ $m$ ][ $n$ ]为灰度连续区域  $m$  的第  $n$  个边界点;  $C(p) = 1$  表示点  $p$  为运动像素
```

```
edgenumber[ $m$ ] = edgenumber[ $m$ ] + 1;
```

```
// 统计灰度连续区域  $m$  的边界点中属于运动区域的点的个数 }
```

```
if(edgenumber[ $m$ ] > Ratio)灰度连续区域  $m$  属于前景运动对象区域}
```

其中,Ratio 是一个经验值,可以通过实验确定,本文 Ratio 值为 20%.这种方法综合利用了空间分割

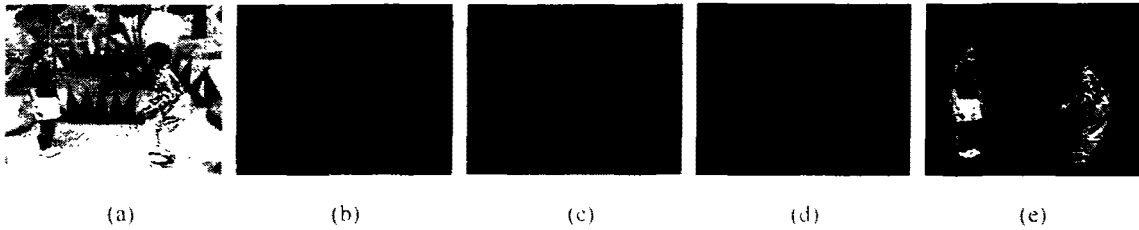
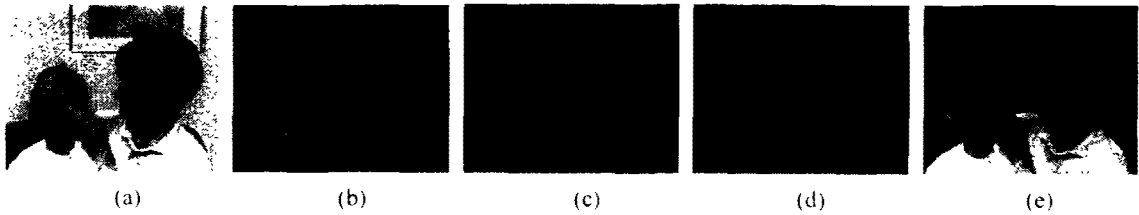
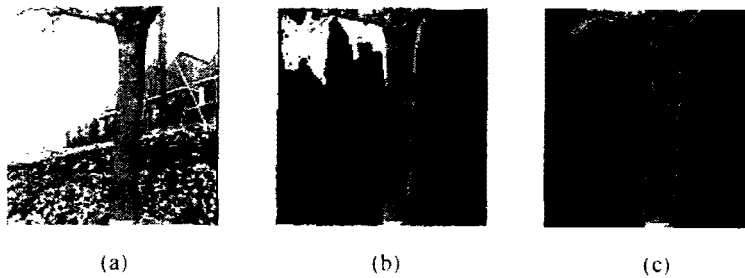
图 5 Children 序列的第 2 帧的分割结果 ($L=1$)Fig. 5 Segmentation of the second frame of "Children" test sequence ($L=1$)图 6 Mother & Daughter 序列的第 15 帧的分割结果 ($L=3$)Fig. 6 Segmentation of the second frame of "Mother & Daughter" test sequence ($L=1$)

图 7 Flowergarden 序列的第 2 帧的分割结果

Fig. 7 Segmentation of the second frame of "Flowergarden" test sequence

信息和像素运动信息,它基于以下事实:(1)一个运动对象区域总是由一些灰度连续区域组成的;(2)当一个灰度连续区域中有较多的像素确定为运动像素时,该灰度连续区域属于运动区域的可能性比较大;(3)一个属于运动对象区域的灰度连续区域,在相邻的两帧图像中可能有很大的重叠区域(如图 2),但是灰度连续区域的边界重叠的比例相对较少.将所有属于运动对象区域的灰度连续区域合并在一起得到运动区域,其余的区域为背景区域.运动区域内的像素标记为 1,背景区域的像素标记为 0,就得到了分割掩模.

用这种方法生成分割掩模的优点是:(1)可以减少一些噪声的影响;(2)可以使分割掩模更完整,一个灰度连续区域只要有超过一定比例(Ratio)的像素被确定为运动像素,则整块灰度连续区域都确定为属于运动区域;(3)分割边界比较精确,这是灰度连续区域的边界对对象区域边界的约束结果.

经过上述几个步骤得到的分割掩模可能还存在一些孤立的噪声区域和边缘毛刺,须做一些后处理.后处理的主要内容是形态滤波,以消除孤立的噪声区域和边缘毛刺.

2 实验结果与分析

采用前面所述的分割方法,分别对 MPEG-4 的测试序列 Mother & Daughter 以及 Children 在 PC 机上进行了分割实验,视频图像为 QCIF 格式.

图 5 和 6 是两组实验结果.其中(a)为原图,(b)为二次帧差图,(c)为灰度连续区域分割结果,图中使用伪彩色显示各个灰度连续区域.根据二次帧差和灰度连续区域分割的结果得到最后的分割掩模,如图(d)和(e)为从原图中分割出来的运动对象区域.在图 6(c)中可以发现,左侧人脸有小部分缺陷,这是由于该处前景区域与背景区域的灰度值非常相似造成的.

从图 5(e)和图 6(e)可以看出,由于采用了二次帧差,并以灰度连续区域分割的边界来限制对象区域的边界,所以较好地抑制了噪声的影响,而且分割区域比较完整,分割区域的边界相当精确.通过图 7 的比较,可以进一步表明基于连续灰度区域分割可以使分割区域更加完整,分割边界更加精确.图 7 (a)为原图,(b)为用文献[6]的方法分割的结果,(c)为用灰度连续区域分割的边界来限制对象区域的结果.

3 结论

本文提出一种基于时空信息的运动视频对象分割方法,综合利用了空间分割信息和像素运动信息,从静止的背景中分割出运动的前景区域.这种方法在分割过程中不需要与人交互,也不需要大量视频帧的统计信息,简化了分割过程.实验结果表明,这种分割方法能较好地抑制噪声的影响,分割的结果比较完整,而且分割的边界相当精确.但是,这种方法只适用于没有全局运动的情况,当存在全局运动的时候,在求帧差之前须先做全局运动补偿,这一部分工作正在进行中.

REFERENCES

- [1] GAO Wen, WU Feng. Researches and developments of MPEG-4 coding. *J. of Computer Research and Development* (高文,吴枫. MPEG-4 编码的现状和研究. 计算机研究与发展), 1999, **36**(6):641—652
- [2] Chuang Gu, Ming-Chieh Lee. Semiautomatic segmentation and tracking of semantic video objects. *IEEE Transactions On Circuits And Systems For Video Technology*, 1998, **8**(5):572—583
- [3] Toklu Candemir, Tekalp A Murat, Erdem Tanju A. Semi-automatic video object segmentation in the presence of occlusion. *IEEE Transactions On Circuits And Systems For Video Technology*, 2000, **10**(4):624—629
- [4] PAN Jin-Hui, LI Shi-Peng, ZHANG Ya-Qin. Automatic extraction of moving objects using multiple features and multiple frames. *IEEE Int. Symposium on Circuits and Systems (ISCAS) 2000*, Geneva, 2000, **1**:36—39
- [5] Neri A, Colonnese S, Russo G. Automatic moving objects and background segmentation by means of higher order statistics. *SPIE 3024*, 1997, **3024**:8—14
- [6] Zhou J Y, Ong E P, Ko C C. Video object segmentation and tracking for content-based coding. *ICME 2000*, 2000, **3**:1555—1558
- [7] Salembier P, Pardas M. Hierarchical morphological segmentation for image sequence coding. *IEEE Transactions On Image Processing*, 1994, **3**(5):629—651
- [8] Chuang Gu, Ming-Chieh Lee. Tracking of multiple semantic video objects for internet applications. *SPIE Conference on Visual Communications and Image Processing '99*, 1999, **3653**:806—820
- [9] Roux L J Le, Schalkwyk J J D Van. An overview of moving object segmentation in video images. *COMSIG 1991 Proceedings*, South African, 1991, 53—57
- [10] ZHAN Jing-Fen, QI Fci-Hu, ZHAO Xue-Chun. Automatic segmentation of moving object and background. *Journal of Infrared and Millimeter Waves* (詹劲峰,戚飞虎,赵雪春. 运动目标和背景的自动分割. 红外与毫米波学报), 1999, **18**(5):343—350