

# 以统计变化检测为基础的实时分割 视频对象新方法

于跃龙 卢焕章

(国防科技大学 ATR 重点实验室,长沙 410073)

**摘要** 为了克服利用变化检测分割视频对象过程中的噪声、复杂运动、暴露背景的影响,提出了一种基于统计变化检测的实时分割视频对象新方法。在该方法中,由于统计变化检测技术是利用  $t$  分布来有效消除噪声的影响,而不需要估计噪声的方差,而且可利用间隔为  $k$  的两帧图像代替连续两帧来进行变化检测,因此可以很好地处理关节运动和慢运动;另外,对两个连续的统计变化检测结果取交集还可以消除暴露背景的影响,并能消除大部分的残留噪声,且几乎不增加计算量,因此统计变化检测可作为视频对象分割的基础,试验结果表明,该方法不仅解决了传统的变化检测过程中的噪声、复杂运动以及暴露背景影响,而且能够自动实时地分割视频对象,以满足 MPEG-4 等基于对象的视频应用。

**关键词** 统计变化检测 视频对象  $t$  分布 MPEG-4

中图分类号: TP391.41 TN941.1 文献标识码: A 文章编号: 1006-8961(2005)01-0098-05

## A New Method for Real-time Segmenting Video Objects Based on Statistical Change Detection

YU Yue-long, LU Huan-zhang

(ATR Key Laboratory, National University of Defense Technology, Changsha 410073)

**Abstract** In order to eliminate the effect caused by noise, complex motion, and uncovered background in the process of real-time segmenting video objects, a new method for real-time segmenting video objects based on statistical change detection is proposed. During the process of statistical change detection, because  $t$ -distribution is used to eliminate the effect of noise, the noise variance is not needed. Two frames at a distance of  $k$  are used for statistical change detection instead of two successive frames, thus articulation motion and slow motion can be better coped with. The intersection between two successive statistical change detection results can eliminate the effect of uncovered background, and most of the residual noise are eliminated at the same time without adding any computation complexity. The experimental results show that the new method solves the problems existing in the process of traditional statistical change detection caused by noise, complex motion, and uncovered background, and it can automatically segment video objects by real-time. It can meet the requirements of many object-based video applications such as MPEG-4.

**Keywords** statistical change detection, video object,  $t$ -distribution, MPEG-4

## 1 引言

新的多媒体数据压缩国际标准 MPEG-4 采用了基于对象的编码技术,并要求把视频序列分解成多个不同的视频对象,然后再对各个视频对象分别编

码,这样一方面可以提高编码效率,另一方面也可以向用户提供存取场景内容的交互功能。视频对象指视频序列中与现实世界物体(如人、汽车等)相对应的区域,如今视频对象已广泛地应用于基于内容的视频检索和高级视频监控等领域,为了成功地进行视频检索和视频监控,首先必须将视频对象从视频

收稿日期:2004-05-13;改回日期:2004-08-11

第一作者简介:于跃龙(1974~),男,2000年获国防科技大学硕士学位,现为国防科技大学 ATR 重点实验室博士生。主要研究方向为图像处理、视频分析、视频压缩。E-mail: yuyuelong\_74@sohu.com; w80210038@sina.com.cn

序列中分割出来。

如今人们已对视频对象的分割技术进行了广泛的研究,其中变化检测技术<sup>[1,2]</sup>计算高效,是一种经常使用的技术。变化检测技术是指通过将当前图像与参考帧图像比较来得到变化区域和未变化区域。由于图像存在噪声,许多背景像素也会被检测为变化像素,所以必须选择合适的门限值来克服噪声的影响。文献[3]提出利用显著性技术来得到门限值,但该方法需要利用噪声的方差,可是一般噪声方差未知,需要进行估计,这样不仅会引进估计误差,同时还将增加计算负担。

帧间变化检测是利用前一帧作为参考帧,并定义变化区域为运动对象和未变化区域为背景。该技术的优点是时间间隔短,且不受光照慢变化的影响,但是存在以下问题:内部纹理细节不丰富的对象,会检测为未变化区域;当对象做非刚性运动(如关节运动)及间歇运动时,停止运动的对象(或者对象部分)则无法检测出来;由于因对象运动而暴露出的背景部分会检测为变化区域,因此可归为对象的一部分。为了消除暴露背景的影响,文献[3]采用在阈值分割后得到的帧差图像的四阶矩图上进行块匹配搜索的方法来估计变化像素的位移,以便判定该像素是否属于视频对象,由于一般搜索过程很耗时,该方法无法实时完成,为此提出了一种基于统计变化检测的实时分割视频对象新方法,该方法不仅有效地解决了阈值选取的问题,而且克服了帧间变化检测的缺点。

## 2 新方法的流程流程

本文提出的基于统计变化检测的实时分割视频对象新方法的流程框图如图 1 所示,其主要包括变化检测和后处理两部分。其中变化检测同其他的变化检测方法比较有以下两点不同:为了克服噪声的影响,根据噪声模型,首先采用 t 分布对帧差图像中以待检测像素为中心的空域窗内的样本值进行显著性检验来确定该像素是否发生了变化;然后,对间隔为  $k$  的两帧图像进行统计变化检测,用以代替帧间变化检测,这样即使对象在连续两帧之间运动较小时,仍然可以分割出完整的视频对象。变化检测后,再对变化检测的结果进行形态学处理,即将分割结果中的变化区域变成封闭区域,并填充由于视频对象纹理不丰富而出现的内部孔洞,最后即得到较为完整的变化区域。

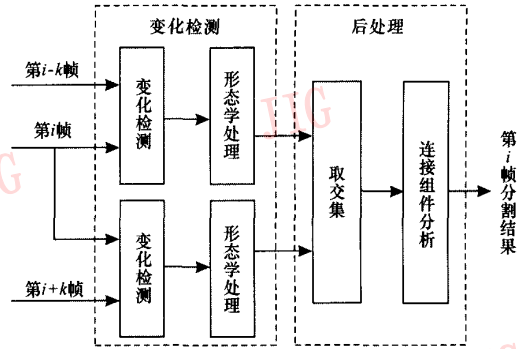


图 1 实时分割视频对象新方法的流程框图

Fig. 1 Flow chart of new method for realtime segmenting video objects

在后处理阶段,本文首先对连续两个变化检测的结果取交集。在第  $i$  帧与第  $i-k$  帧之间的变化检测结果中,将发生变化的区域记为  $M_{i,i-k}$ ,其中  $i$  为当前帧的序数。 $M_{i,i-k}$  包含第  $i$  帧中的对象和因对象的运动而暴露出的背景区域 UB (Uncovered Background), UB 位于变化区域内与运动方向相反的一侧,或者认为  $M_{i,i-k}$  包含第  $i-k$  帧中的对象和由于对象的运动而遮挡的背景区域 CB (Covered Background), CB 位于变化区域内与运动方向相同的一侧。通过分析可知,连续两个变化检测中的变化区域  $M_{i,i-k}$  和  $M_{i,i+k}$  都包含第  $i$  帧的视频对象,此外,  $M_{i,i-k}$  还包含 UB 区域,  $M_{i,i+k}$  还包含 CB 区域, UB 区域和 CB 区域在位置上分别位于视频对象的两侧,将  $M_{i,i-k}$  和  $M_{i,i+k}$  取交集后就可以得到视频对象区域,同时消除了 UB 区域和 CB 区域。对于变化检测过程中未完全消除的噪声像素点,由于这些像素点是随机的,不具有时域连续性,因此取交集可以进一步消除噪声的影响。接着对取交集的结果采用连接组件分析<sup>[3]</sup>去除残留的噪声,由于连接组件分析在图像中用于标识每个单独的区域和计算区域的面积,且只将小于一定门限值的区域作为噪声忽略,因此在清除掉小面积区域的同时,并不破坏对象的形状信息。

## 3 实时分割视频对象新方法的实现

本文假设相机固定,当相机运动时,首先进行全局运动补偿,然后采用类似的分割方法进行分割。

### 3.1 基于 t 分布的统计变化检测

由于像素值的变化是由对象运动或者噪声引

起,因此变化检测过程中需要克服噪声的影响。由于噪声主要来源于相机,且可以将相机噪声建模为加性 Gaussian 噪声<sup>[4]</sup>,其方差记为  $\delta_{\text{camera}}^2$ ,据此序列的第  $i$  帧图像可以表示为

$$f(x, y, i) = \hat{f}(x, y, i) + n(x, y, i) \quad (1)$$

其中,  $f(x, y, i)$ 、 $\hat{f}(x, y, i)$ 、 $n(x, y, i)$  分别表示第  $i$  帧图像的观测值、实际图像像素值和噪声。根据式(1),第  $i$  帧与参考帧的帧差表示为

$$\begin{aligned} f_{\text{dif}}(x, y, i) &= f(x, y, i) - f_{\text{reference}}(x, y) \\ &= \hat{f}(x, y, i) - \hat{f}_{\text{reference}}(x, y) + \\ &\quad n(x, y, i) - n_{\text{reference}}(x, y) \end{aligned} \quad (2)$$

其中,  $f_{\text{reference}}(x, y)$  表示参考帧图像像素值,  $n_{\text{reference}}(x, y)$  表示参考帧图像中的噪声,  $f_{\text{dif}}(x, y, i)$  表示帧差像素值。令  $d(x, y, i) = n(x, y, i) - n_{\text{reference}}(x, y)$ , 可以认为  $n(x, y, i)$  和  $n_{\text{reference}}(x, y)$  是概率密度函数相同和独立的随机变量,由概率论知识可知,  $d(x, y, i)$  是加性零均值 Gaussian 随机变量,方差  $\delta^2 = 2\delta_{\text{camera}}^2$ , 即

$$p(d) = \frac{1}{\sqrt{2\pi\delta^2}} e^{-\frac{d(x, y, i)^2}{2\delta^2}} \quad (3)$$

下面利用显著性技术来检测帧差图像中的每个像素点是否发生变化。一般来讲,由于背景点的小邻域窗内的像素属于背景点的概率较高,因此在以待检测像素为中心的邻域窗内对该像素进行检验的可靠性优于对单个像素的检验。设第  $i$  帧视频以  $(x, y)$  为中心的邻域窗  $W$  像素  $W_{(x, y, i)}$  为待检测像素,窗的大小为  $q = (2N + 1)(2M + 1)$ ,  $q$  是窗内的像素数,  $(2N + 1)$  和  $(2M + 1)$  分别是窗的宽和高。由于帧差图像中各像素点的噪声是互相独立的,如果窗内所有不为零的帧差值都是由噪声引起的,则这些值的均值  $\mu$  为零,所以,可进行如下假设:  $H_0: \mu = 0, H_1: \mu \neq 0$ 。

根据数理统计知识,在噪声方差未知的情况下,对均值进行假设检验应采用  $t$  分布,并根据窗内的样本构造如下统计量

$$t = \frac{\bar{f}_{\text{dif}}(x, y, i)}{(s(x, y, i) / \sqrt{q})} \quad (4)$$

其中,

$$\begin{aligned} \bar{f}_{\text{dif}}(x, y, i) &= \frac{1}{q} \sum_{l=-N}^N \sum_{m=-M}^M |f_{\text{dif}}(x+l, y+m, i)| \\ s(x, y, i) &= \sqrt{\frac{1}{q-1} \sum_{l=-N}^N \sum_{m=-M}^M (f_{\text{dif}}(x+l, y+m, i) - \bar{f}_{\text{dif}}(x, y, i))^2} \end{aligned}$$

假设给定的显著性水平为  $\alpha$  (一般  $\alpha$  的取值为  $10^{-1} \sim 10^{-3}$ ), 则可以得到如下拒绝域:

$$|t| \geq t_{\frac{\alpha}{2}}(q-1) \quad (5)$$

其中,  $t_{\frac{\alpha}{2}}(q-1)$  的值可以通过查  $t$  分布表得到。若满足式(5)的像素点为发生变化的像素点,则这些点的集合就构成变化区域  $M_{i, \text{reference}}$ 。变化检测掩膜值  $c_{i, \text{reference}}(x, y)$  可以表示为

$$c_{i, \text{reference}}(x, y) = \begin{cases} 1 & (x, y) \in M_{i, \text{reference}} \\ 0 & \text{其他} \end{cases} \quad (6)$$

### 3.2 视频对象的分割

对于视频对象第  $i$  帧,在进行视频对象分割时,首先利用 3.1 节介绍的统计变化检测方法,将  $i-k$  帧作为参考帧进行变化检测,即得到变化区域  $M_{i, i-k}$ 。  $M_{i, i-k}$  包含视频对象区域  $O_i$  和由视频对象的运动而暴露出的背景区域  $R_{\text{UB}}$  (下角 UB 为 uncovered background 缩写), 即

$$M_{i, i-k} = O_i \cup R_{\text{UB}} \quad (7)$$

同样,以第  $i+k$  帧为参考帧,通过与第  $i$  帧进行统计变化检测来得到变化区域  $M_{i, i+k}$ ,  $M_{i, i+k}$  包含视频对象区域  $O_i$  和由于视频对象的运动而覆盖的背景区域  $R_{\text{CB}}$  (下角 CB 为 covered background 缩写), 即

$$M_{i, i+k} = O_i \cup R_{\text{CB}} \quad (8)$$

由于对象向同一方向运动,且  $R_{\text{UB}}$  和  $R_{\text{CB}}$  分别位于视频对象的两侧方向,所以

$$R_{\text{UB}} \cap R_{\text{CB}} = \emptyset \quad (9)$$

然后,将  $M_{i, i-k}$  和  $M_{i, i+k}$  取交集即可得到分割的视频对象,即

$$M_{i, i-k} \cap M_{i, i+k} = (O_i \cup R_{\text{UB}}) \cap (O_i \cup R_{\text{CB}}) = O_i \quad (10)$$

最后进行取交集运算就可以消除  $M_{i, i-k}$  中包含的 UB 区域。同时,由于假设检验存在一定的虚警概率,因此在  $M_{i, i-k}$  中包含一些受噪声影响而发生变化的背景像素,另外,由于噪声是随机的,不具有时域连续性,所以取交集后可以进一步消除大部分的残留噪声。如果对得到的结果再进行连接组件分析,就可以得到最终的视频对象掩膜图像。

## 4 实验结果分析

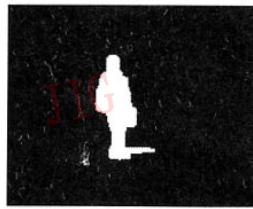
为验证本文方法是否有效,采用多个视频序列进行了试验,图 2 和图 3 是对 MPEG-4 测试序列 Hall object 和 Akiyo 的分割结果,其中 Hall object 视频序列包含的视频对象做关节运动,而 Akiyo 序列

包含的视频对象做慢运动。为了解本方法优点,还将本文方法的分割结果与帧间变化检测结果进行了比较(图 2,图 3)。一般  $k$  的取值主要取决于视频对象的运动速度,如果视频对象的运动速度较快,则  $k$  可以取较小的值,反之, $k$  需要取较大值。如果本文方法用于嵌入式系统,则  $k$  的取值还要和存储量结合起来考虑,因为增大  $k$  值,需要增加系统的存储量,所以本文此处不考虑存储器的影响,由于 Hall object 和 Akiyo 序列中的视频对象运动速度不快,所以试验中  $k$  值取 5,窗的大小为  $5 \times 5$ ,显著性水平

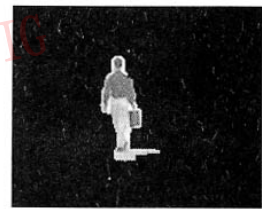
$\alpha = 0.01$ 。膨胀腐蚀选取的基元大小为  $3 \times 3$ ,连接组件分析使用 8-连通,最终的结果中区域面积小于 16 的区域认为是噪声而清除掉。图 2(a)、图 2(d)分别为 Hall object 视频序列第 57、58 帧;图 2(b)是利用本文的算法将 Hall object 视频序列第 52、57 帧的变化检测掩膜图像与 Hall object 视频序列第 57、62 帧的变化检测掩膜图像取交集后,再进行连接组件分析得到的分割掩膜图像;图 2(c)是根据图 2(a)和图 2(b)得到的分割结果;图 2(e)是根据图 2(a)和图 2(b)得到的分割结果;图 2(f)是 Hall object 视频序列第 57、58 帧进行帧间变化检测后再进行形



(a) Hall object 视频序列第 57 帧



(b) 本文算法分割掩膜图像



(c) 本文算法分割结果



(d) Hall object 视频序列第 58 帧



(e) 帧间变化分割掩膜图像



(f) 帧间变化分割结果

图 2 Hall object 视频的分割结果

Fig. 2 Segmentation results of video sequence Hall object



(a) Akiyo 视频序列第 82 帧



(b) 本文算法分割掩膜图像



(c) 本文算法分割结果



(d) Akiyo 视频序列第 83 帧



(e) 帧间变化分割掩膜图像



(f) 帧间变化分割结果

图 3 Akiyo 视频的分割结果

Fig. 3 Segmentation results of video sequence Akiyo

态学处理得到的掩膜图像;图 2 (f) 是帧间变化检测的最终分割结果。由于视频对象做关节运动,因此帧间变化检测不能分割出完整的视频对象,而使用本文的方法则可以很好地分割视频对象。图 3 是 Akiyo 头肩视频序列的分割结果,其连续两帧之间的运动非常小。通过试验结果可以看出,由于本文的分割方法可以很好地处理慢运动,并较好地克服了帧间变化检测技术存在的问题,因而比帧间变化检测的分割结果有了显著的提高,而且本文的分割方法计算复杂度低,可以实现以 25 帧/s 的速度分割视频序列。

## 5 结论及进一步的工作

本文提出了一种新的基于统计变化检测的实时分割视频对象新方法。试验结果表明,该方法不仅很好地解决了自动阈值选取问题,而且克服了帧间变化检测存在的缺点,可以较为准确地分割视频对象。

由于使用该方法分割视频对象得到的分割结果中包含阴影区域,而且当多个视频对象之间发生遮挡时,该方法效果不好,因此如何消除阴影区域的影响,以及在多个对象发生遮挡时,如何进行有效的处理,是下一步的研究重点。

### 参考文献 (Reference)

- 1 Neri A, Colonnese S, Russo G, *et al.* Automatic moving object and background separation [J]. *Signal Processing*, (Special Issue on Video Sequence Segmentation for Content-based Processing and Manipulation, Elsevier), 1998, **66**(2): 219 ~ 232.
- 2 Cavallaro A, Ebrahimi T. Accurate video object segmentation through change detection [A]. In: *Proceedings of IEEE International Conference on Multimedia and Expo [C]*, Lausanne, Switzerland, 2002: 26 ~ 29.
- 3 Haralick R, Shapiro L. *Computer and Robot Vision vol. 1 [M]*. Reading, Mass, USA: Addison-Wesley, 1992.
- 4 Aach Til, Kaup André, Mester Rudolf. Statistical model-based change detection in moving video [J]. *Signal Processing*, 1993, **31**(2): 165 ~ 180.