

运动目标的自动空时分割算法

李仁杰 余松煜

(上海交通大学图像通信所, 上海 200240)

摘要 在视频应用中,运动目标的提取是一个重要的研究课题。为了对运动目标进行更有效的分割,提出了一种从视频序列中自动提取运动目标的空时分割算法。该算法在时域分割中采用基于齐异矢量消除的目标检测方法来获得运动目标的初始模板。通常,该初始模板具有不连续的边界和一些“孔”。为了得到较为完整的目标区域,用具有距离约束的区域生长算法来补偿初始模板。而在空域分割中,分水岭分割则通过考虑全局信息来增强其分割的精确性。然后,精确的运动目标即可通过空时融合模块提取出来。试验结果表明,该空时分割算法是有效的。

关键词 空时分割 区域生长 分水岭 齐异矢量

中图分类号: TN911.73 文献标识码: A 文章编号: 1006-8961(2007)10-1931-04

Unsupervised Spatio-temporal Segmentation of Moving Objects in Video Sequences

LI Ren-jie, YU Song-yu

(Institute of Image Communication and Information Processing, Shanghai Jiaotong University, Shanghai 200240)

Abstract The extraction of moving objects is an important and fundamental research topic for many video applications. This paper addresses an unsupervised spatial-temporal segmentation scheme to extract moving objects from video sequences. In temporal segmentation, an outlier rejection (OR) based object detection approach is proposed to extract initial temporal masks, followed by region growing with a distance constraint to compensate initial temporal masks accurately in order to link discontinuous boundaries and fill some holes. In spatial segmentation, watershed segmentation considering the global information improves the accuracy of segmentation in the spatial domain. By using a fusion module, moving objects are extracted. Experiments on various sequences have successfully demonstrated the validity of the proposed scheme.

Keywords segmentation, region growing, watershed, outlier

1 引言

在许多视频应用中,运动目标的检测是一个基本且重要的研究课题。目前已提出许多的算法用来解决该问题。

在运动目标的分割算法中,基于空时分割的算法是一类重要的方法。该类方法利用了视频的空时关系,通过融合空域和时域的分割结果,得到了较好的分割结果^[1]。本文提出了一种新的空时分割算法,且该算法在分割过程中不需要用户的干预。通

常,时域分割中的困难主要是如何精确地估计全局运动矢量。为了解决这个问题,本文提出了一种基于齐异点消除的全局矢量估计算法。为了检测运动目标,先将每一运动场分解为不重叠的块;然后通过比较每一块的运动矢量和全局运动矢量来得到目标的初始模板。该方法虽增强了时域分割对噪声的鲁棒性,但得到的初始模板通常有不连续的边界和一些“孔”。为了便于后续分割,这些初始模板需要被补偿。本文用具有距离约束的区域生长算法来实现补偿,其增加的距离约束项可防止生长点偏离于相应的目标区域。由于通常的分水岭分割算法没有考

基金项目:国家自然科学基金项目(60502034)

收稿日期:2007-06-05; 改回日期:2007-07-25

第一作者简介:李仁杰(1977~),男,现为上海交通大学图像通信所博士研究生。研究方向为图像和视频处理。E-mail: lrjie@sjtu.edu.cn

虑全局信息,而通过考虑全局信息,则可以增强空间分割的精确性,为此本文采用了权重梯度图像的方法来增强分水岭分割的精确性。

2 时空分割算法

图 1 描述了所提算法的总体框架。

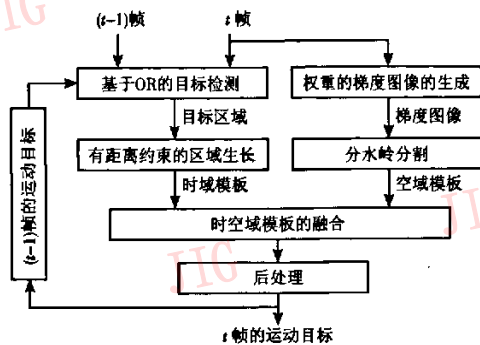


图 1 所提出的时空分割算法的总体框架
Fig. 1 Proposed segmentation scheme

2.1 时域模板的提取

大家知道,视频中像素的运动包括全局运动和局部运动。其中全局运动由摄像机的运动引起;局部运动由运动目标产生。由于通常期望的运动目标的运动不同于背景的全局运动,因此,粗糙的目标可通过分析运动场得到。设 $I_t(x, y)$ 为视频中第 t 帧坐标为 (x, y) 点的像素值。对于连续帧 $\{I_t(x, y)\}$ 和 $\{I_{t-1}(x, y)\}$, 可先用块匹配算法^[2]来计算 t 帧的前向运动场,其匹配块的尺寸为 5×5 ; 然后,对运动场进行模板尺寸为 3×3 的中值滤波^[3]。中值滤波可除去一些噪声矢量。

设 $v_t = (v_t^x, v_t^y)$ 为得到的运动矢量,其中 v_t^x 和 v_t^y 分别为运动矢量在 x 轴和 y 轴方向的分量。在计算全局运动矢量时,与全局运动矢量差别较大的奇异矢量应该被消除。由于运动目标的运动矢量通常不同于全局运动矢量,因此目标上的每个点的运动矢量应该标记为奇异矢量,其不能用于全局运动矢量的计算。全局运动矢量的计算包括以下两步:

(1) 消除由运动目标产生的奇异矢量,这个过程主要通过把在 $t-1$ 帧检测到的运动目标投射到第 t 帧来实现的。设 $R_{t-1}(i)$ 为在 $t-1$ 帧的目标模板 M_{t-1} 中的第 i 个目标区域, $N_{t-1}(i)$ 是 $R_{t-1}(i)$ 的面积(像素的个数)。 $R_{t-1}(i)$ 上的点 (x, y) 可据式

(1) 投射。

$$\begin{cases} \hat{x} = x + \vec{v}_{t-1}^x(i) \\ \hat{y} = y + \vec{v}_{t-1}^y(i) \end{cases} \quad (1)$$

其中, (\hat{x}, \hat{y}) 为预测坐标,而 $R_{t-1}(i)$ 的运动矢量 $(\vec{v}_{t-1}^x(i), \vec{v}_{t-1}^y(i))$ 则通过下式计算:

$$\begin{cases} \vec{v}_{t-1}^x(i) = \left(\sum_{(x,y) \in R_{t-1}(i)} v_{t-1}^x(x,y) \right) / N_{t-1}(i) \\ \vec{v}_{t-1}^y(i) = \left(\sum_{(x,y) \in R_{t-1}(i)} v_{t-1}^y(x,y) \right) / N_{t-1}(i) \end{cases} \quad (2)$$

处理完所有的目标区域后,即可得到预测的二进制模板 \hat{M}_{t-1} 。这样就可根据 \hat{M}_{t-1} 中的预测坐标,将 t 帧运动场中相应点的运动矢量标记为奇异矢量,图 2 给出了投射方法的示意图。

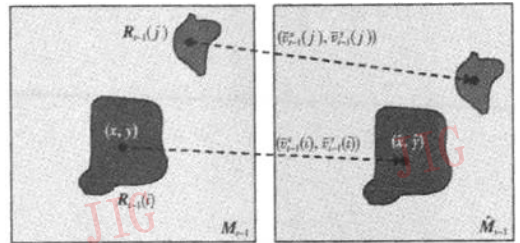


图 2 投射方法示意图

Fig. 2 Illustration of the projection method

(2) 第 t 帧的全局运动矢量 $\vec{v}_t = (\vec{v}_t^x, \vec{v}_t^y)$ 可通过不断去除残余的奇异矢量的迭代算法得到。假设经过 $n-1$ 次迭代,去除奇异矢量后的运动场的平均运动矢量为 $\vec{v}_t^{(n-1)}$ 。在 n 次迭代,若一个未被标记的矢量 $v_t(x, y)$ 和 $\vec{v}_t^{(n-1)}$ 间的距离 $d(v_t(x, y), \vec{v}_t^{(n-1)})$ 大于阈值 T_d , 则将该矢量标记为奇异矢量。而矢量 $u = (u^x, u^y)$ 和 $v = (v^x, v^y)$ 间的距离 $d(u, v)$ 被定义为

$$d(u, v) = \max(|u^x - v^x|, |u^y - v^y|) \quad (3)$$

当连续两次迭代的平均运动矢量不再变化或者迭代次数 n 达到最大值 K 时,则迭代过程停止。最终得到的平均矢量即为全局矢量 \vec{v}_t 。 T_d 和 K 的经验值分别设为 0.7 和 4。注意 $|\vec{v}_t^x|$ 或 $|\vec{v}_t^y|$ 的值可能小于 1。

为了鲁棒地检测到运动目标,将当前的运动场分解为 $M \times N$ 个不重叠的块,块的大小为 $B \times B$ 。块 (m, n) 根据式(4)决定是否为目标块。

$$J_t(m, n) = \begin{cases} 1 & \text{if } d(\vec{v}_t(m, n), \vec{v}_t) \geq T_{\text{motion}} \\ 0 & \text{其他} \end{cases} \quad (4)$$

其中, $J_i(m, n)$ 为块 (m, n) 的二进制的值, $\bar{v}_i(m, n)$ 是块 (m, n) 中所有的运动矢量的平均值。如果 $J_i(m, n) = 1$, 则该块中的所有像素被置为 1; 否则为 0。这样即可得到目标的初始模板 $F_i^{initial} = \{F_i^{initial}(x, y)\}$ 。阈值 T_{motion} 的范围为 $0.5 \leq T_{motion} \leq 1.6$ 。

为了补偿粗糙的目标模板, 这里提出一种新的区域生长算法, 即首先标记模板 $F_i^{initial}$ 中的每一个 8 连通目标区域(点值为 1)。相同目标区域的点的符号相同, 而不同区域的点的符号则不同。假设 p 为一个生长种子点。区域 $S(p)$ 生长于该种子 p 。定义一个点 $q = (k, l)$ 和模板 $F_i^{initial}$ 中被标记的一个目标区域 $Q_i(i)$ 之间的距离为

$$\hat{d}(q, Q_i(i)) = \arg \min_{(x, y) \in Q_i(i)} (\sqrt{(x-k)^2 + (y-l)^2}) \quad (5)$$

量度 $\delta(q, S(p))$ 反映了点 q 的灰度与区域 $S(p)$ 的平均灰度 $\bar{I}_i(S(p))$ 间的差别程度。其被定义为

$$\delta(q, S(p)) = |I_i(q) - \bar{I}_i(S(p))| \quad (6)$$

基于 sequentially-sorted list (SSL)^[4], 区域生长算法的具体实现过程如下:

(1) 从初始目标模板中选择初始种子, 如果一个被标记点的 8 个邻域点不全被标记, 则该点为种子点;

(2) 从初始种子中选择一个未曾处理的种子 a , 其属于第 i 个目标区域 $Q_i(i)$, 如果所有的初始种子已被处理, 则生长停止; 否则继续步骤(3);

(3) 如果 a 的所有 8 个邻域点已被标记, 则返回到步骤(2), 否则把没有被标记的邻域点放在 SSL, 继续步骤(4);

(4) 如果 SSL 为空, 返回到步骤(2), 否则, 继续步骤(5);

(5) 从 SSL 中移出第 1 个点 b , 计算相应的距离 $\hat{d}(b, Q_i(i))$ 和量度 $\delta(b, Q_i(i))$ 。如果 $\hat{d}(b, Q_i(i))$ 小于阈值 T_d 且 $\delta(b, Q_i(i))$ 小于阈值 T_s , 则继续步骤(6); 否则返回到步骤(4);

(6) 用 a 的符号标记 b , 并把 b 的 8 个邻域点中没有标记的点放入 SSL。如果新放入的点已存在于 SSL 中, 则删除新加入的点, 更新区域 $S(b)$ 的平均灰度, 返回到步骤(4)。

生长停止后, 如果一个目标区域的面积小于阈值 T_A , 则将该区域合并到背景区域中; 然后将所有被标记的点合并, 并置为 1; 其他点置为 0。这样就可得到最终的时域运动模板 $F_i = \{F_i(x, y)\}$ 。

2.2 权重梯度图像的分水岭分割

空间分割包括权重梯度图像的产生和分水岭分割^[5]两步。其中权重梯度图像的目的是用全局信息增强期望的运动目标的边界。

本文应用 Sobel 梯度算子计算初始的梯度(边界)图像 $G_i = \{G_i(x, y)\}$ 。一般地, 与期望目标的一个边界点所连通的边界点的数量要大于其他点(如区域内部点)所连通的边界点的数量。这样即可通过下式来对图像 G_i 的每个边界点进行加权。

$$\hat{G}_i(x, y) = (A_i(x, y) / (\bar{A}_i(x, y) + 1)) G_i(x, y) \quad (7)$$

其中, $A_i(x, y)$ 是与点 (x, y) 连通的边界区域的面积, $\bar{A}_i(x, y)$ 为图像 G_i 中所有的边界区域的平均面积。均匀量化器^[6]把实值的 $\hat{G}_i(x, y)$ 量化成 0~255。这样即可得到量化后的图像; 然后用分水岭算法分割该量化图像, 即可得到具有更为精确边界的空间区域。

2.3 空时分割的融合

对于空间分割结果, 其每个区域反映了灰度和运动的一致性。假设一个空间区域与运动目标的时域模板 F_i 中目标区域相交的那部分区域所占该空间区域的百分率为 r , 如果 r 大于阈值 T_{union} , 则该空间区域为目标区域; 否则为背景区域。处理完所有的空间区域后, 即可得到第 t 帧的目标模板 M_t 。

3 试验结果

本文试验采用文献[1]和文献[7]作为参考算法, 取阈值 $T_d = 2.5$ 和 $T_s = 6$ 。这两个阈值应该足够小, 以避免不期望的点生长到目标区域中。本文取 $T_A = 150$ 。

图 3 给出了 QCIF 格式的 Coastguard 序列的分割结果。文献[1]算法的融合阈值为 0.6。这里, $T_{motion} = 0.6, T_{union} = 0.7, B = 8$ 。从图 3 可见, 本文算法的分割结果是比较完整的, 且包含较少的噪声区域。文献[1]算法的分割结果有较多的噪声区域。文献[7]算法的分割结果中, 目标包含了一些不期望的区域, 且部分区域未被检测出。主观上看, 本文算法的分割结果要好于其他两种算法的分割结果。另外, 文献[7]算法采用了精确的初始目标模板。

图 4 给出了 CIF 格式的 Stefan 序列的连续分割结果。这里, $T_{motion} = 0.8, T_{union} = 0.55, B = 16$ 。文献[7]算法的初始模板采用了本文算法的分割结果的第 2 帧。图 5 给出了尺寸为 200×200 的 Street 序列的分割结果。

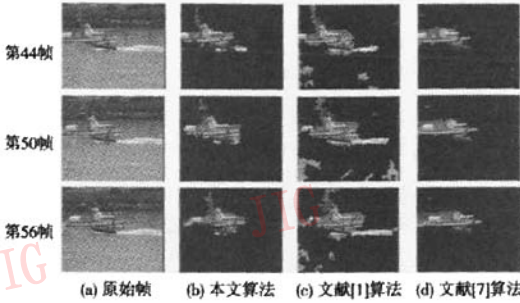


图 3 Coastguard 序列的分割结果

Fig. 3 Results with the three algorithms

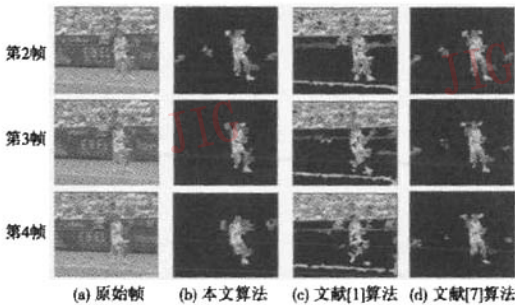


图 4 Stefan 序列的分割结果

Fig. 4 Results with the three algorithms

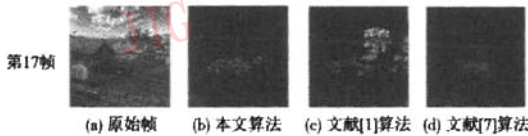


图 5 Street 序列的分割结果

Fig. 5 Results with the three algorithms

为了客观地比较分割结果,表 1 对比了 3 个算法的平均分割误差量度^[8]。从表 1 可见,本文算法具有较小的分割误差。

表 1 平均分割误差量度比较

Tab. 1 Mean segmentation error measures

算法	不同序列的平均分割误差量度		
	Coastguard 序列	Stefan 序列	Street 序列
本文算法	0.41	0.90	0.63
文献[1]算法	1.40	6.40	3.52
文献[7]算法	0.52	0.92	0.49

4 结 论

本文提出了一种自动地提取运动目标的时空分割算法。该算法无需用户提供初始目标模板,即可成功地从复杂的场景中分割出运动目标,而且它还增强了分割的精确性和鲁棒性。试验结果表明,该算法是有效的。未来的工作主要为如何精确地估计运动场。

参考文献 (References)

- Kim M, Choi J G. A VOP generation tool: Automatic segmentation of moving objects in image sequences based on spatial-temporal information[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(8): 1216 - 1226.
- Ezhilarasan M, Thambidurai P. An efficient block matching algorithm for fast motion estimation in video compression[A]. In: Proceedings of International Conference on Signal Processing & Communication [C], Bangalore, India, 2004.
- Pitas I, Venetsanopou A. Nonlinear Digital Filters: Principles and Application[M]. Norwell, MA, USA: Kluwer, 1990: 154 - 156.
- Adams R, Bischof L. Seeded region growing[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(6): 641 - 647.
- Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(6): 583 - 598.
- Gray R M, Neuhoff D L. Quantization[J]. IEEE Transactions on Information Theory, 1998, 44(6): 2325 - 2382.
- Xia J, Wang Y. A spatio-temporal video analysis system for object segmentation[A]. In: Proceedings of the third Symposium on Image and Signal Processing and Analysis [C], Rome, Italy, 2003: 18 - 20.
- Wollborn M, Mech R. Refined procedure for objective evaluation of video object generation algorithm[S]. Doc. ISO/IEC JTC1/SC29/WG11 M3448, Mar. 1998.