

## 多模型鲁棒估计的运动分割

任建峰<sup>1</sup>, 郭雷<sup>1</sup>, 沈玉利<sup>1,2</sup>

(1. 西北工业大学自动化学院, 陕西西安 710072; 2. 湛江海洋大学电子工程系, 广东湛江 524025)  
(rjff@163.com)

**摘要:**提出了一种基于多模型结构鲁棒估计的运动分割算法。首先对视频处理对象进行基于四叉树的分裂合并, 获取鲁棒估计的初始运动数目以及相应的运动模型的初始参数, 然后通过参数估计, 不断更新模型参数, 之后通过把每个运动区域和几个运动模型相关联, 来同时估计多个运动的区域, 最后通过小物体的运动检测方法检测出小的运动物体, 最终达到分割的目的。试验证明该算法取得了比较明显的结果。

**关键词:**视频对象分割; 鲁棒估计; 四叉树的分裂合并

**中图分类号:** TP391.41 **文献标识码:** A

## Motion segmentation based on multi-model robust estimation

REN Jian-feng<sup>1</sup>, GUO Lei<sup>1</sup>, SHEN Yu-li<sup>1,2</sup>

(1. School of Automation, North Western Polytechnic University, Xi'an Shaanxi 710072, China;  
2. Department of Electrical Engineering, Zhanjiang Ocean University, Zhanjiang Guangdong 524025, China)

**Abstract:** A novel motion segmentation algorithm based on multi-model robust estimation was proposed. Firstly, the split and merge technique based on quad-tree was proposed to obtain the initial motion numbers of robust estimation and corresponding initial parameters of motion model. Then the parameters were updated by using estimating parameters, and the motion models were merged based on the motion consistency. Lastly small motion objects could be detected by using small-object detection techniques. Experimental results show that the method can obtain promising results.

**Key words:** video object segmentation; robust estimation; split and merge based on quad-tree.

### 0 引言

视频分割就是要研究如何把视频序列分解成一系列的运动物体, 同时还要跟踪这些运动物体的运动。本文首先定义两个概念: 运动区域和运动物体。运动区域为具有一致运动的区域, 而且这个区域可能是运动物体的一个部分; 运动物体为由一致空间特征的相邻运动区域的合并。因此, 物体是由一个或多个区域组成的。一个刚性的运动物体只由一个运动区域组成, 但是一个非刚性的物体就可能由多个运动的区域组成。

本文提出了一个基于多模型结构鲁棒估计的运动分割算法来分割运动的区域。通过基于四叉树的分裂合并的方法给出了运动物体数目的初始估计以及这些模型的初始运动参数, 然后通过把每个运动区域和几个运动模型相关联, 来同时估计多个运动的区域。最后通过小物体的运动检测方法检测出小的运动物体, 最终达到较满意的分割效果。

### 1 多模型鲁棒估计的运动分割算法

#### 1.1 多模型估计

在鲁棒估计模型中, 主要利用参数以及残余误差来建立运动矢量模型, 利用残余误差区分出格点和非出格点, 同时减少出格点的影响。本文提出的多模型鲁棒估计能够同时估计多个运动, 主要通过关联不同的运动区域和相应的运动模型, 减少了伪出格点的数目, 降低了污染度。

假设图像  $D$  可以表示成  $K$  个独立的区域  $M$  和一个呈正态分布的加性噪声  $\eta$  之和。

$$D = \sum_{k=1}^K s^{(k)} M^{(k)}(X(K)) + \eta \quad (1)$$

其中加性噪声  $\eta$  服从零均值的正态分布,  $M^{(k)}(X(K))$  表示为参数  $X(K)$  的模型,  $s^{(k)}$  为支撑模板。其中每个模型  $M^{(k)}(X(K))$  都是对估计图像建立的参数模型。在运动分割中, 不同的运动可以利用具有不同参数的仿射运动模型表示。由于仿射运动模型是线性函数, 因此可以利用  $M$  估计来计算运动分割。在公式(1)条件下, 计算不同模型的残余误差:

$$r(i, j) = d(i, j) - \sum_{k=1}^K s^{(k)} M(i, j; X^{(k)}) \quad (2)$$

参数的估计可以利用公式(3)的等式:

$$\{X^{(k)}\} = \arg \min_{\{X^{(k)}\}} \sum \rho(r(i, j)) \quad (3)$$

从公式(2)、(3)可以看出, 为了获得最优参数  $\{X^{(k)}\}$ , 则需要知道最优分割  $s^{(k)}$ 。这个就是著名的“鸡生鸡蛋—鸡蛋生鸡”问题, 可以贝叶斯估计的方法解决。但是这种统计问题的求解需要用到全局优化, 需要大量的计算, 因此很少采用。

为了解决这个问题, 本文提出了一个方法, 该方法的特点就是把图像表示成树状结构的数据, 树叶表示图像的一个区域。如果由树叶表示的区域具有一致性, 那么这些区域就称为非出格点区域。如果具有非一致性, 那么这些区域就称为出格点区域。来自不同模型的数据组成的出格点, 在后面的估计过程中将被抛弃。利用单一的模型来拟合区域, 目标函数可以写

成:

$$X = \arg \min_{(i,j) \in R} \rho(r(i,j), X) = \arg \min \Delta(R, X) \quad (4)$$

其中:  $r(i,j) = d(i,j) - M(i,j; X)$

对每个区域进行鲁棒估计可能会造成模型数目的过估计,而且对于噪声也敏感。为了加强算法的鲁棒性,在所有的非出格点区域上进行鲁棒估计。一个区域的  $\Delta$  最小化必须考虑所有的非出格点区域。为了考虑所有的非出格点区域以及寻找最佳的  $K$  个模型,引进了  $s_r^{(k)}$  作为树结构中区域一致的支撑模板。假定模型的数目是已知的,  $s_r^{(k)}$  可以通过以公式(5)获得:

$$s_r^{(k)} = \begin{cases} 1 & \Delta(R, x^{(k)}) = \min \{ \Delta(R, x^{(i)}) \}, i = 1, 2, \dots, K \\ 0 & \text{否则} \end{cases} \quad (5)$$

结合  $s_r^{(k)}$  可以获得最优模型参数:

1) 初始化参数  $X^{(1)}, X^{(2)}, \dots, X^{(K)}$ :

$$\begin{cases} X^{(1)} = \arg \min \sum_R s_r^{(1)} \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(1)}))}{(d_{i,j} - y_{i,j}(X^{(1)}))} \\ X^{(2)} = \arg \min \sum_R s_r^{(2)} \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(2)}))}{(d_{i,j} - y_{i,j}(X^{(2)}))} \\ \vdots & \\ X^{(K)} = \arg \min \sum_R s_r^{(K)} \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(K)}))}{(d_{i,j} - y_{i,j}(X^{(K)}))} \end{cases} \quad (7)$$

2) 更新每个区域的  $\Delta^{(k)}$ :

$$\begin{cases} \Delta^{(1)} = \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X^{(1)}))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(1)}))}{(d_{i,j} - y_{i,j}(X^{(1)}))} \\ \Delta^{(2)} = \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X^{(2)}))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(2)}))}{(d_{i,j} - y_{i,j}(X^{(2)}))} \\ \vdots & \\ \Delta^{(K)} = \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X^{(K)}))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(K)}))}{(d_{i,j} - y_{i,j}(X^{(K)}))} \end{cases} \quad (8)$$

3) 更新每个区域的  $s_r^{(k)}$ :

$$s_r^{(k)} = \begin{cases} 1 & \Delta^{(k)} = \min \{ \Delta^{(1)}, \Delta^{(2)}, \dots, \Delta^{(K)} \} \\ 0 & \text{否则} \end{cases} \quad (9)$$

4) 重新更新参数  $X^{(1)}, X^{(2)}, \dots, X^{(K)}$ :

$$\begin{cases} X^{(1)} = \arg \min \sum_R s_r^{(1)} \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(1)}))}{(d_{i,j} - y_{i,j}(X^{(1)}))} \\ X^{(2)} = \arg \min \sum_R s_r^{(2)} \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(2)}))}{(d_{i,j} - y_{i,j}(X^{(2)}))} \\ \vdots & \\ X^{(K)} = \arg \min \sum_R s_r^{(K)} \sum_{(i,j) \in R} w_{i,j} (d_{i,j} - y_{i,j}(X))^2 & \text{其中 } w_{i,j} = \frac{\Psi(d_{i,j} - y_{i,j}(X^{(K)}))}{(d_{i,j} - y_{i,j}(X^{(K)}))} \end{cases} \quad (10)$$

5) 重复 2) ~ 4), 直到参数  $X^{(1)}, X^{(2)}, \dots, X^{(K)}$  不再变化。

### 1.4 区域标记

区域标记的目的是进一步细化基于块的图像分割结果,这主要通过把每个像素和估计过程中的参数运动模型结合起来实现的。而且,把具有相似运动的模型合并起来。

基于块的分割主要把一个块分为单一运动/主运动的块,或者含多个运动边界的多运动块。但是这种方法并不适用于最终的分割结果,因为这些块中含有运动边界,通常不能很好地描述运动一致的区域。为了解决这个问题,提出了一个结合模型合并的区域标记方法用于标记运动一致的区域。为了

$$\{X^{(k)}\} = \arg \min \sum s_r^{(k)} \Delta(R_r^n, X^{(k)}) \quad (6)$$

为了求解公式(6),必须知道模型数目的个数以及  $s_r^{(k)}$ , 然后使用迭代的方法。参数更新的过程,模型数目以及  $s_r^{(k)}$  的初始估计是通过四叉树的分裂合并完成的。

### 1.2 基于四叉树分裂合并的参数初始化

为了消除出格点的影响,本文采用了基于四叉树的分裂合并策略。基于四叉树的分裂合并方法能够给出主运动数目的初始估计。区域分裂主要根据时域信息,如果一个区域具有非时域一致性,或非一致的运动,那么就会被不断的分裂成四个更小的区域,直到划分的区域达到运动一致性,或达到区域大小的下界。

### 1.3 运动参数的估计

在获得了运动数目的初始估计以及运动模型的参数初始估计之后,本文采用递归的方法来求解公式(6)。

关联运动区域和相应的运动模型,在计算相应的运动模型的残余误差基础上标记每个像素。而且这种标记可以分为三种:确定的标记、模糊的标记以及未指定的标记。一个确定的标记是指像素的运动可以用一个特定的参数运动模型表示。一个模糊的标记指像素的运动可以用多于一个的运动模型表示,而未指定的标记是指像素的运动不能用任何一个估计模型表示。

通过对残差图像  $R$  进行中值滤波得到中值残差  $\bar{r}_{i,j}$ :

$$\bar{r}_{i,j} = \text{median\_filter}(R)$$

$$R = \{r_{i,j} | i = 1, \dots, M; j = 1, \dots, N\} \quad (11)$$

如果只有一个残余误差小于给定的阈值  $T_{residual}$ , 这个点就标记为确定的点; 如果多余一个残余误差小于给定的阈值  $T_{residual}$ , 就标记为模糊的点; 如果没有一个残余误差小于给定的阈值  $T_{residual}$ , 就标记为未标记的点。

其中具有确定标记的点具有相同的运动, 而具有模糊标记的点被认为含有不同的运动模型。这些点为比较运动模型的相似性提供了重要的基础。关于未标记的点将在下面的运动物体检测中给予进一步处理。

### 1.5 小运动物体的检测

下面还必须对未标记的点进行处理, 因为其中可能含有小的运动物体。这些未标记的点主要由于两个原因造成的: 出格点或小运动物体。出格点和小运动物体的区别在于出格点分布不规则, 而小运动物体具有内在的一致性特征。因此, 本文采用数学形态消除出格点。首先, 对图像中未标记的点标记 1, 其余的标记为 0, 这样产生一幅二进制图像  $I$ , 然后对这个二进制图像采用开运算, 消除不规则的出格点。完成数学形态学滤波后, 就可以在图像中发现一个或多个区域, 这些区域就是小的运动物体。

## 2 实验结果分析

本文利用了乒乓球比赛序列作为实验对象。图 1(a)、(b) 给出了视频序列的连续两帧, 用于光流的计算。图 1(c) 给出了计算得到的光流。可以看出, 由于光流计算的不精确性, 所以得到的光流数据存在噪声干扰。图像中存在三种不同的运动物体: 乒乓球、球拍和人。由于乒乓球和球拍是刚性物体, 所以可以使用一个仿射模型表示。而人非刚性物体, 不可能只用一个仿射模型表示, 必须使用多个运动模型表示。

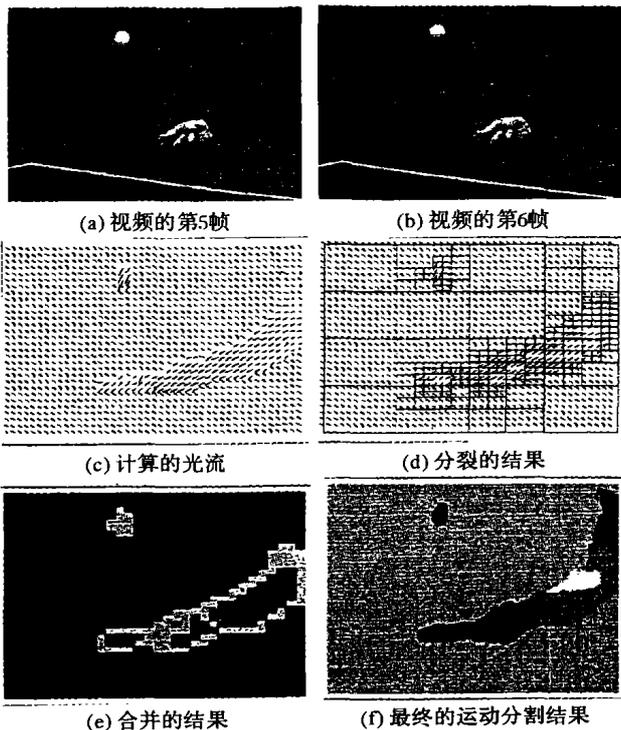


图 1 乒乓球视频的运动分割实验结果

分裂合并的初始阈值设置为 6.0, 图 1(d) 给出了基于四叉树的分裂结果, 合并的结果如图 1(e)。初始的时候, 包括背景的运动一共存在七个仿射运动, 通过区域标记后, 有两个合并到一起。图 1(f) 给出了最终的分割结果。可以看出物体乒乓球、球拍和人都能被检测出来。其中人是由四个分割

区域组成, 因为非刚性物体不能用一个模型描述, 必须使用多个运动模型表示。另外, 通过分裂合并的方法不能检测出乒乓球, 最后通过小物体的运动检测才能检测出来。

## 3 结语

由于传统的鲁棒估计, 不但对于噪声比较敏感, 而且在多运动的检测中, 容易受到伪出格点的影响。因此提出一个多模型鲁棒估计的运动分割算法用于解决多运动的检测, 利用多个运动模型同时对多个运动建立模型, 通过把一个区域和多个运动模型相关联。

本算法首先对图像进行基于四叉树的分裂合并算法, 给出了鲁棒估计的初始运动数目以及相应的运动模型的初始参数。然后通过参数估计, 不断更新模型的参数, 最终达到较满意的分割效果。

### 参考文献:

- [1] HAMPEL FR, ROUSSEEUW PJ, RONCHETTI E, et al. Robust Statistics: the approach based on influence functions[M]. New York: John Wiley & Sons, 1986.
- [2] STEWART C. Bias in Robust Estimation Caused by discontinuities and Multiple structures[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(8).
- [3] MEIER T, NGAN KN. Automatic Segmentation of moving objects for video object plane generation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(5): 525 - 538.
- [4] DARRELL T, PENTLAND A. Cooperative Robust estimation using layers of support[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(5): 474 - 487.
- [5] HUANG Y, PALANIAPPAN K, ZHUANG X. Optic flow field segmentation and motion estimation using a robust genetic partitioning algorithm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(12): 1177 - 1189.

## 2005 年网络、通信与信息系统学术会议征文通知

2005 年专委会的学术年会将于 2005 年 9 月下旬在西南交通大学召开。有关征文事宜通知如下:

一、征文内容(主要包括但不限于以下方面):

- ①下一代互联网(NGI)研究与实践;
- ②软件教育与软件产业发展;
- ③无线通信与接入技术;
- ④网络计算: 网格计算、普适计算、对等计算;
- ⑤网络与信息安全;
- ⑥电子政务、电子商务;
- ⑦数据挖掘。

二、投稿要求:

通过会议学术组评审被接收的论文将在《计算机科学》2005 年 9 月增刊上发表; 稿件字数限 5000 字, 论文格式请按《计算机科学》投稿要求撰写(详细要求请咨询《计算机科学》编辑部, 电话: (023) 63500828, Email: jsjcx@swic.ac.cn)。

稿件可采用邮寄打印件(一式 2 份)加软盘邮寄方式, 也可采用打印件(一式 2 份)加电子邮件方式, 凡截稿期内缺少电子版或打印稿者, 恕不审稿。

投稿截止时间: 2005 年 4 月 30 日

录用通知时间: 2005 年 5 月 15 日

三、投稿请寄: (610031) 四川成都 西南交通大学, 计算机与通信工程学院 窦 军(副教授)

E-mail: doujun@tom.com

中国计算机学会西南学会

网络与信息系统专业委员会

四川省计算机学会

2005 年 1 月 15 日