

一种基于区域收缩的运动分割算法

李智慧 黄凤岗

(哈尔滨工程大学计算机科学与技术学院, 哈尔滨 150001)

摘要 运动分割需要估计出每个运动的运动模型参数和运动支持区。为准确地确定运动支持区,在最大后验边缘概率(MPM-MAP)算法的基础上,提出了一种新的基于区域收缩的运动分割算法,用于确定运动支持区。该算法先以属于某种运动概率最大的像素为备选像素,然后通过区域收缩选定备选像素密度最大的区域为支持区范围。此外,还提出了一种新的运动模型参数初值的估计方法,并将差分图像包围盒的确定和区域收缩相结合,用于运动模型参数的初值估计,该方法先估计运动区域,再估计运动模型参数,并通过运动分解、合并和消亡来获得准确的运动个数。实验结果表明,该方法是有效的。

关键词 运动分割 最大后验边缘概率 区域收缩

中图法分类号:TN919.8 文献标识码:A 文章编号:1006-8961(2008)08-1496-05

A Motion Segmentation Algorithm Based on Region Shrinking

LI Zhi-hui, HUANG Feng-gang

(School of Computer Science and Technology, Harbin Engineering University, Harbin 150001)

Abstract Motion segmentation needs to estimate the model parameters of every motion as well as its supporting region. On the basis of maximum posterior marginal probability (MPM-MAP) algorithm this paper presents a new algorithm based on region shrinking to locate the supporting area. In this algorithm pixels of maximum probabilities belonging to a motion are considered to be candidate pixels for supporting region. Then the region shrinking algorithm is used to determine the region of maximum density of the candidate pixels to be the range of supporting area. Moreover, this paper presents a new approach combining the bounding box's defining with region shrinking to estimate the initial parameters of motions. By motion dividing, motion incorporation and motion elimination the accurate number of motions can be obtained. The results of experiments show the validity of this method.

Keywords motion segmentation, MPM-MAP, region shrinking

1 引言

众所周知,运动图像序列分割是计算机视觉中的一个重要问题,而解决该问题的难点在于既要估计出运动模型的参数,又要估计出每个运动所对应的区域(即支持区)。文献[1]提出将运动区域用层来描述,这样一幅运动图像就由若干层构成,每个层是一个运动的支持区,而所估计的运动则一般用仿射模型表示。

用于运动模型估计的一类较为有效的方法是基于概率框架的方法。文献[2]~文献[6]则是采用EM算法估计运动模型参数,但是不包括空间一致性约束。文献[7]~文献[10]则采用了基于马尔可夫随机场(MRF)的一致性约束,用于去除估计出运动场的噪声。文献[11]提出了最大后验边缘概率(MPM-MAP)运动分割算法,该算法分两步估计运动模型参数和支持区,不仅将MRF一致性约束结合到概率公式中,而且应用了文献[12]提出的一种MRF的快速算法。

收稿日期:2006-12-19; 改回日期:2007-04-09

第一作者简介:李智慧(1966~),女,副教授。2001年于哈尔滨工程大学获得硕士学位。研究方向为图像处理与计算机视觉。

E-mail: lizhh@126.com

目前所采用的一致性方法主要是 MRF 方法和一些光流算法中常采用的局部平滑方法。MRF 方法局部平滑的效果好一些,但计算复杂度高。虽然 MRF 方法可以在一定程度上去除孤立点和小空洞,但遇到稍大的孤立区域和空洞就无能为力了。在运动估计中,由于在目标中央一些颜色一致的区域中很容易产生较大的空洞,所以采用 MRF 并不能从根本上解决空间一致性问题。

采用概率框架的方法一般要假设初值,包括运动目标个数和每个运动模型参数的初值,这就使得运算复杂度增加,而且经常使运动估计落入局部最小值而影响运动估计的准确性,因有时这些局部最小值的误差很大。

本文提出了一种新的通过区域收缩来确定支持区的方法,该方法可以对运动区域(通过 MPM-MAP 算法估计)精确定位。不同于一般的局部一致性方法,该方法选取以均值为中心的标准差范围内一个完整的区域作为运动支持区,其好处是可以获得一个完整的区域,而忽略其中中心的孔洞和外部的孤立小区域,并且可得到较鲁棒的运动区域的定位,这对下一步获得准确的运动模型参数估计极有帮助。区域收缩算法不仅可以用在运动模型参数的估计上,还可以用在差分图像上。本文先通过求取差分图像的包围盒(包围某区域的矩形范围),然后用区域收缩法定位,进而获得支持区的运动模型参数的初值,即先获得支持区,再估计运动模型参数的初值。这样不仅增加了运动模型参数初值的准确性,并使运算速度加快,结果更加准确。

2 基于贝叶斯框架的 MPM-MAP 算法

解决运动分割较为有效的一类方法是基于概率框架的模型拟合方法,即首先定义若干参数未知的运动模型,然后通过最大后验概率准则(MAP)确定其参数。本文采用文献[11]提出的概率框架,即通过以下两个参数定义每种运动,其中一个为运动模型参数 θ (即仿射运动的参数),另一个是支持区参数 b (其是一个开关指示函数或二值图像,当像素 r 属于第 k 个运动的支持区 R_k 时, $b_k(r)$ 为 1, 否则为 0),最大后验概率问题可以转化为最大似然函数问题。似然函数的定义为

$$Q(b, \theta) = \log(p(f, b | \theta)) = Q_1 + Q_2 \quad (1)$$

$$Q_1 = \sum_{k=1}^K \sum_{r \in R_k} b_k(r) \log(h_k(r))$$

$$h_k(r) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left[-\frac{(f_2(\Phi(r; \theta_k)) - f_1(r))^2}{2\sigma_k^2}\right]$$

$$Q_2 = \sum_{k=1}^K \sum_{r \in R_k} \frac{\eta_k}{S_{R_k}}$$

上式中, f_1 和 f_2 为图像序列中相邻的两帧, r 为位置向量, $\{\Phi(r; \theta_k) (k=1, 2, \dots, K)\}$ 为一个仿射运动模型集合,每个模型以 θ_k 为参数, $\Phi(r; \theta_k)$ 所对应的支持区为 R_k 。 Q_1 为数据逼近项(同文献[11]); Q_2 为本文提出的支持区平滑项定义,即为支持区中属于该运动的像素的密度。分割时,以属于某种运动概率最大的像素为备选像素, $\sum \eta_k$ 为 R_k 中的备选像素个数, S_{R_k} 为 R_k 的面积, η_k 和 S_{R_k} 中隐含了支持区参数 b 。这里 b 与 θ 是未知数,可通过下式得到

$$\begin{aligned} (\hat{b}, \hat{\theta}) &= \operatorname{argmax}_{b, \theta} (p(f, b | \theta)) \\ &= \operatorname{argmax}_{b, \theta} (Q(b, \theta)) \end{aligned} \quad (2)$$

理想情况下,可根据 MAP 法则,按式(1)和式(2)求与最大后验概率对应的参数 b 和 θ 。但是由于式(2)中的两项不容易同时迭代,计算代价太高,所以本文采用 MPM-MAP 估计算法,即假设 b 已知,首先求使似然函数 $Q(b, \theta)$ 最大的 θ ; 然后假设 θ 已知,再求出使 $Q(b, \theta)$ 最大的 b 值,即可求出支持区。这相当于求对于 θ 和 b 的最大边缘概率。算法步骤如下:

- (1) 在差分图像上采用第 5 节介绍的方法估计运动的个数与运动初始区域 R_k , 设定 θ 的初值,并根据 θ 的初值和 b , 求使 Q 值最大的 θ 值,令 $i=1$;
- (2) 根据 θ , 采用区域收缩方法求 b 值;
- (3) 已知 b , 根据式(1)和式(2)计算最佳的 θ , 作为 $\theta^{(i+1)}$;
- (4) 重复步骤(2)、(3)直到迭代收敛,即 Q 值大于某个阈值 K_0 ;
- (5) 通过边缘跟踪方法确定准确的运动区域。

有多种确定使 Q 值最大的运动模型参数 θ 的方法,本文采用文献[13]中的过松弛算法来确定,该算法属于一种梯度下降算法,其优点是可以保证收敛。

3 支持区的确定——区域收缩算法

当运动模型参数 θ 确定以后,还需要确定每个像素的 b 值,即确定每个运动所对应的区域。确定对应区域时,首先根据式(3)确定 R_k 的备选像素集 \hat{R}_k , 一个备选像素是在 K 个支持区中属于 R_k 的概率最大的像素。

$$\hat{R}_k = \{r | h_k(r) > \hat{h}_k(r), \hat{k} \in \{1, \dots, K\} \text{ 且 } \hat{k} \neq k\} \quad (3)$$

如图 1 所示,其中所有白色像素就为属于运动 1 的备选像素。

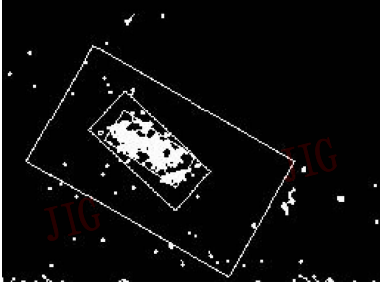


图 1 区域收缩过程

Fig. 1 Process of region shrinking

本文将平滑约束定义为支持区中属于该运动区域的像素的密度(见式(1))。假设 R_k 中的备选像素的概率相同,则式(2)可通过式(4)逼近。

$$(\eta_k, S_{R_k}) = \operatorname{argmax} Q_2 \quad (4)$$

式(4)的含义是求出备选像素密度最大的区域为支持区。对于任何分布,备选像素密度最大的区域一定是在以备选像素坐标的均值 μ 为中心的一定范围区域中,均值 μ 可以根据备选像素密度最大区域的像素坐标计算出来。根据切比雪夫不等式,随机变量在以均值 μ 为中心的 3σ (标准差)范围内的概率大于 0.89;在 2σ 范围内的概率大于 0.75。假设每个备选像素的概率相同,且 σ 范围内的概率大于 0.5,则较大密度的备选像素的范围可以通过均值和方差确定,并可通过一种区域收缩的方法来确定合适的支持区。

在考虑理论概率和实际情况的同时,本文选择 σ 为保留区域的范围值,选择矩形为保留区域的形状,并根据均匀分布来确定具体参数。计算时的变量值为像素坐标,根据均匀分布,每个像素的概率为像素个数的倒数。区域收缩算法中的参数可通过计算备选像素坐标的均值 μ 和协方差矩阵 C 得到。可通过对矩阵 C 做相似对角化来求出矩阵特征值 λ_1, λ_2 和特征向量 v_1, v_2 。 μ 为矩形区域的中心; v_1, v_2 指示出矩形区域两中轴的方向;矩形区域的两边的长度分别为 $2l_x$ 和 $2l_y$,其可由式(5)确定,即

$$l_x = \frac{1}{2} \sqrt{12 \cdot \lambda_1} \quad l_y = \frac{1}{2} \sqrt{12 \cdot \lambda_2} \quad (5)$$

区域收缩算法(Matlab 伪代码)如下:

```
function [ $\mu, v_1, v_2, l_x, l_y$ ] = Region_shrink( $B, k_2$ );
%  $B$ : 备选像素集, (二值图像);  $k_2$ : 密度阈值
```

```
%  $\mu, v_1, v_2, l_x, l_y$ : 均值, 特征向量和矩形区域的两半
边长
```

```
 $D = B$ ;
```

```
 $d = \text{Density}(D)$ ; %  $d$  为矩形区域  $D$  的密度
```

```
while  $d < k_2$ 
```

```
     $\mu = \text{Mean}(D)$ ; % 求矩形区域  $D$  中备选像素的均值
```

```
     $C = \text{Co\_var}(D)$ ; % 求矩形区域  $D$  中备选像素的协方差
    矩阵
```

```
    [ $V, L$ ] = eig( $C$ ); % 求  $C$  的特征向量( $V$ )和特征值( $L$ )
```

```
     $v_1 = V(1, :)$ ; % 特征向量
```

```
     $v_2 = V(2, :)$ ;
```

```
     $l_x = \text{sqrt}(L(1, 1) * 12) / 2$ ;
```

```
     $l_y = \text{sqrt}(L(2, 2) * 12) / 2$ ; % 根据特征值计算的矩形区
    域的两个半边长
```

```
     $D = \text{Keep}(D, \mu, v_1, v_2, l_x, l_y)$ ; % 保留由( $\mu, v_1, v_2, l_x,$ 
 $l_y$ )确定的矩形区域内的像素, 矩形区域外的所有点去掉
```

```
     $d = \text{Density}(D)$ ;
```

```
end
```

其中 eig 和 sqrt 为 Matlab 函数;“%”以后为注释内容。

4 初始参数的估计和准确形状的获取

本文通过差分图像来估计运动个数及运动支持区初值,由于初值估计得较为准确,所以经过较少的迭代次数即收敛,而且分割结果较好。

4.1 用包围盒确定运动所在区域

包围盒为包围一个连通区域的最小矩形。通过差分图像中一个连通区域的包围盒,就可以确定一个初始运动区域。由于直接将每个包围盒当作一个运动区域会使运动个数过多、运动区域太小而失去准确性,所以可以通过若干次合并,使属于同一目标的包围盒合并为一个较大的区域。但合并次数太多,又可能使多个目标处于同一个区域而加大估计复杂性。本文采用一次合并结果。

4.2 用区域收缩算法确定初始支持区

由包围盒所确定的运动区域比较大,不够准确,但可以在包围盒内,采用区域收缩算法来确定较准确的支持区初值。因为差分图像像素较少,所以区域收缩算法仅一次迭代便可确定支持区。

4.3 根据支持区确定运动模型参数

先根据支持区参数 b 的初值来假设一个运动模型参数的初值,然后采用过松弛算法进行估计。由于一般的算法都对初值敏感,所以本文对每种运动假设了 4 个方向的运动模型参数,即 4 个初值,对每

个初值分别进行迭代,取 Q 值最大的结果作为运动模型参数的估计结果。这种做法虽然使初值计算次数增多,但却增加了初值的准确性,从而使后面的迭代次数明显减少。

4.4 错误结果处理

由于实际情况中会出现估计运动个数多于或少于实际数量的情况,所以需要对可能出现的错误进行处理。针对可能出现的情况,本文采用了以下 3 种方法进行处理。

(1) 运动合成

当估计出的两个运动模型参数其距离小于一个阈值 ε 时,则认为两种运动为一种,即将两种运动合并。

(2) 运动消亡

当区域收缩算法在规定次数 n_i 内无法收敛于一个较大的 Q_2 值(例如 k_2)时,而且该运动区域较小(小于 $T_{length} \times T_{length}$),则该区域可能并非运动区域,即删除该运动模型参数,使运动消亡。

(3) 运动分解

当出现与运动消亡同样的情况,而且该运动区域较大(大于 $T_{length} \times T_{length}$),则该区域可能不仅仅包含一种运动,此时需检查原来是否有包围盒合并。如果有,则撤消原来的合并操作,重新估计运动;如果没有,则假设该区域运动数为 2,并采用文献[11]方法进行估计。如果误差仍然很大,则增加运动个数,重新估计,直到误差小于阈值为止。

4.5 获取准确的运动区域

区域收缩所获得的仅是区域的位置,而且通过收缩后的矩形区域获得准确的运动区域是比较简单的一步,其可以通过很多方法实现,其中活动轮廓方法是较准确的一种,但较费时。本文先用形态学闭操作对区域边缘进行平滑,然后采用简单的轮廓跟踪获取准确的运动目标区域。

5 实验结果

本文通过一些实例证明了该方法的有效性。图 2~图 5 显示了对 Hamburg Taxi 图像(第 25、26 帧,如图 2 所示)进行处理的结果。实验时,首先通过差分图像确定 6 个包围盒(图 3(a)),再经过合并(合并结果如图 3(b)所示)估计出 5 个支持区(图 3(c)),并分别假设运动位移初值为 $(\pm 2, \pm 2)$,再进行过松弛迭代,估计出的运动初值如表 1 所示;然后采用 MPM-MAP 算法估计运动场及运动区域。

由于初值估计得较为准确,致使整个 MPM-MAP 算法的迭代次数仅为 2 次,运动 1 在第 1 次迭代时就得到最佳运动模型参数,运动 2 和 3 在第 2 次得到最佳运动模型参数。运动 4 和 5 在第 1 次迭代时,分别被运动 2 和 3 合并。最终得到的运动模型参数如表 2 所示。

表 1 运动模型参数初值

Tab. 1 Initial parameters of motion models

运动	θ_0	θ_1	θ_2	θ_3	θ_4	θ_5
1	-7.89	0.94	-0.035	-1	0	1
2	-6	1	0	-1	0	1
3	6	1	0	1	0	1
4	5	1	0	4	0	1
5	6	1	0	2	0	1

表 2 运动模型参数的最终结果

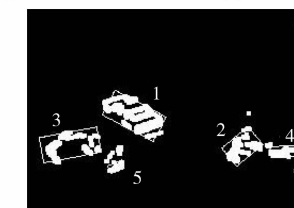
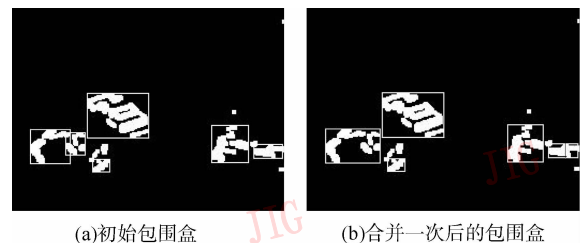
Tab. 2 Final parameters of motion model

运动	θ_0	θ_1	θ_2	θ_3	θ_4	θ_5
1	-7.1	0.94	-0.035	-1	0	1
2	-5	1	0	-1	0	1
3	-3	0.99	0.07	1	0	1



图 2 Hamburg Taxi 图像

Fig. 2 Hamburg Taxi images



(c)区域收缩算法定位的运动支持区

图 3 运动支持区初值估计过程

Fig. 3 Estimation process of initial supporting regions

图 4(a) 显示了对运动 1 采用区域收缩算法得到的初始分割结果。当最后的运动模型参数和运动支持区确定后,则需要采用形态学算子对初始分割结果进行平滑,其平滑结果如图 4(b) 所示。最后经轮廓跟踪得到的分割结果及在原图像中的效果如图 5 所示。



(a)获得的支持区

(b)经平滑后的结果

图 4 运动 1 的一次迭代过程

Fig. 4 One iteration of the first motion



(a)分割结果

(b)在原图像中的效果

图 5 运动分割结果

Fig. 5 Results of motion segmentation

与文献[10]的 EM 算法相比,本文的算法在执行时间上远远少于文献[10]算法。本文算法的平均执行时间大约为文献[10]算法的 20%。文献[10]方法是采用均值场逼近 MRF,虽然比其他方法(如模拟退火等算法)速度快,但执行时间仍然较长。相比较而言,文献[11]算法的执行时间较短,但其分割结果要依赖于初值。如果采用文献[11]方法,用本文假设的运动模型参数初值则不可能收敛到正确解,因为该假设值与真实值差距较大,难以估计出正确的支持区。从中可以看出,本文方法鲁棒性较好。本文算法在其他实例上也取得了较好效果。

6 结 论

本文在 MPM-MAP 算法的基础上,采用一种新的基于收缩支持区的算法来确定运动支持区。此外,本文还提出一种新的运动参数初值估计方法,将差分图像包围盒的确定和区域收缩相结合用于运动初值估计,并通过运动分解、合并和消亡来获得准确的运动数。

由于本文方法能够估计出运动的中心区域,所以

可以较准确地估计出运动模型参数,进而可获得较好的分割结果。而且,由于它不会因初值假设得不准确而使迭代终止于误差较大的解,因而鲁棒性较强。

本文的方法对那些外部轮廓较简单的运动目标很有效,如运动的人体、汽车等。但是对于像自然景物,如树木这样形状复杂的目标仍效果不好,这是今后应该解决的问题。

参考文献 (References)

- 1 Wang J Y A, Adelson E H. Representing moving images with layers[J]. IEEE Transactions on Image Processing, 1994, 3(5): 625 ~ 638.
- 2 Weiss Y. Smoothness in layers: motion segmentation using nonparametric mixture estimation [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], San Juan, Puerto Rico, 1997: 520 ~ 527.
- 3 Weiss Y, Adelson E H. A unified mixture framework for motion segmentation: incorporating spatial coherence and estimating the number of model[A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition[C], Vienna, Austria, 1996: 321 ~ 326.
- 4 Elias D P, Kingsbury N G. The recovery of a near optimal layer representation for an entire image sequences[A]. In: Proceedings of IEEE Conference on Image Processing [C], Washington, DC, USA, 1997: 735 ~ 738.
- 5 Vasconcelos N, Lippman A. Empirical Bayesian motion segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(2): 217 ~ 221.
- 6 Zhang W, Fang X. Moving vehicles segmentation based on Bayesian framework for Gaussian motion model [J]. Pattern Recognition Letters, 2006, 27(9): 956 ~ 967.
- 7 Zhang J. The mean field theory in EM procedures for Markov random fields[J]. IEEE Transactions on Signal Process, 1992, 40(10): 2570 ~ 2583.
- 8 Zhang J. The mean field theory in EM procedures for Markov random fields[J]. IEEE Transactions on Image Processing, 1993, 40(10): 2570 ~ 2583.
- 9 Jaakkola T S, Jordan M I. Improving the mean field approximation via the use of mixture distributions[A]. In: Jordan M I. (Ed.), Learning in Graphical Models [C], Dordrecht: MA, USA: Kluwer Academic Publishers, 1998: 163 ~ 173.
- 10 Celeux G, Forbes F. EM procedures using mean field-like approximations for Markov model-based image segmentation [J]. Pattern Recognition, 2003, 36(1): 131 ~ 144.
- 11 Calderon F, Marroquin J L. The MPM-MAP algorithm for motion segmentation [J]. Computer Vision and Image Understanding, 2004, 95(2): 165 ~ 183
- 12 Marroquin J L, Velasco F, Rivera M, Nakamura M. Gauss-markov measure field model for low-level vision[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(4): 337 ~ 348.
- 13 Black M, Anandan P. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields[J]. Computer Vision and Image Understanding, 1996, 63(1): 75 ~ 104.