

文章编号 1004-924X(2014)12-3409-10

## 结合分段复合权值与多策略的视觉运动目标跟踪

苏延召\*, 李艾华, 王 涛, 张 玮, 金广智  
(第二炮兵工程大学 502 教研室, 陕西 西安 710025)

**摘要:** 由于视觉监控中运动目标跟踪的准确性易受遮挡、摄像机运动、目标外观变化等因素的影响, 本文提出了一种结合分段复合权值与多策略的视觉跟踪算法。该算法首先利用目标、背景以及候选区域特征信息建立分段的复合权值得到目标的位置概率分布。然后结合空间一致性和滞后阈值分割目标位置概率图以进一步抑制噪声干扰, 同时通过分析分段复合权值变化判断目标遮挡, 调整目标跟踪候选范围, 并结合目标历史尺度信息对当前目标尺度进行自适应调整。最后, 对目标以及背景区域信息进行动态更新以适应目标外观与场景变化。与典型算法进行的对比实验结果表明: 该算法能够有效地应对目标遮挡与摄像机运动等因素的影响, 实验时对各组视频的平均处理时间约为 10 ms 左右, 适用于复杂场景条件下运动目标的实时跟踪。

**关键词:** 视觉跟踪; 运动目标; 复合权值; 多策略; 均值漂移

**中图分类号:** TP391 **文献标识码:** A **doi:** 10.3788/OPE.20142212.3409

## Visual tracking of moving objects based on piecewise fusion weight and multi-strategy

SU Yan-zhao\*, LI Ai-hua, WANG Tao, ZHANG Wei, JIN Guang-zhi

(502 Faculty, The Second Artillery Engineering University, Xi'an 710025, China)

\* Corresponding author, E-mail: syzllh@163.com

**Abstract:** As the accuracy of moving object tracking in video surveillance is disturbed by occlusion, camera moving and target appearance changing, an algorithm based on piecewise fusion weight and multi-strategy was proposed. Firstly, the piecewise fusion weight was constructed by combining the feature of object, background and candidate regions to obtain the likelihood image of object location. Then, the likelihood image was segmented with the spatial coherent and hysteresis threshold to suppress noise interference. Meanwhile, the object occlusion was determined and handled by analyzing the change of the piecewise fusion weight and enlarging the candidate area. Furthermore, the object scale was adaptively adjusted according to history and current scales. Finally, object information and background regions were dynamically updated to adapt to the object appearance and scene changing. Experimental results compared with other traditional methods show that the proposed algorithm is applicable to process the moving object tracking in low-contrast scenes in real time, and the average processing time for different video images is 10 ms, which means that the algorithm is suitable for the

收稿日期: 2014-08-04; 修订日期: 2014-10-01.

基金项目: 国家自然科学基金青年基金资助项目(61201449)

moving object tracking in complex scenes.

**Key words:** visual tracking; moving object; fusion weight; multi-strategy; mean shift

## 1 引言

运动目标跟踪是视觉监控、人机交互、智能机器人等计算机视觉领域的一项重要研究内容<sup>[1]</sup>。视觉跟踪的目的是要在视频序列中准确地定位出感兴趣的运动目标。然而由于场景与目标之间的相对变化,跟踪问题常常面临光照变化、场景遮挡、目标外观变化等诸多因素的挑战。因此如何克服这些挑战,提高跟踪算法的准确性、实时性、鲁棒性成为目前研究的焦点与难点<sup>[2]</sup>。

多年来国内外学者对视觉跟踪问题进行了深入研究,提出了许多高效的跟踪算法,这些算法大部分可以划分为两种类型:基于均值漂移(Mean Shift)<sup>[3-4]</sup>框架和基于滤波理论(如卡尔曼滤波<sup>[5]</sup>、粒子滤波<sup>[6-8]</sup>等)框架。由于 Mean Shift 算法具有较高的准确性与良好的实时性,因而受到广泛关注。Comaniciu 等人基于核密度颜色直方图对目标外观进行建模,利用 Mean Shift 算法对候选区域的概率分布进行迭代搜索,直到找到与目标分布最相似的区域,从而实现对目标的跟踪<sup>[9]</sup>。Collins 等人提出利用方差比对 RGB 空间组合形成的特征进行在线选择,再利用这些特征对目标区域进行区分,生成目标位置概率图,最后采用 Mean Shift 算法对位置概率图进行搜索,该方法有效地提高了跟踪算法的准确性<sup>[10]</sup>。Birchfield 等人提出采用二阶空间颜色直方图表征目标的外观,并推导了对应的 Mean Shift 跟踪算法,通过加入空间信息提高了目标外观模型的稳定性,该方法对于目标的姿态变化具有较好的适应性<sup>[11]</sup>。Jeyakar 等人通过对目标区域进行分块,并结合边缘信息应对局部遮挡、光照变化等因素对运动目标跟踪的影响,取得了较好的效果,但算法的运算量较高<sup>[12]</sup>。文献<sup>[13-14]</sup>结合目标区域周围的背景信息对目标颜色直方图进行加权,得到了更为准确的目标权重图,从而提高了跟踪算法抗遮挡与漂移的能力。Wang 等人将局部背景区域、候选区域、目标区域信息相结合,提出了一种复合权值计算方法,进一步增强了目标权重图的准确性,但其跟踪结果容易受到噪声的干扰<sup>[15]</sup>。Choi 等

人结合直方图级与像素级的模板匹配来增强跟踪的准确性与可靠性,算法利用核跟踪产生的权值图将模板匹配转化为一个优化问题,实验结果表明了算法的有效性<sup>[16]</sup>。文献<sup>[17-18]</sup>利用嵌入式平台对图像跟踪算法进行设计与实验验证,结果表明系统满足实时性要求。

本文以分段复合权值表征目标位置概率分布图,通过融合阈值分割、遮挡处理、尺度自适应策略的 Mean Shift 迭代以及动态更新实现复杂场景下的视觉运动目标跟踪。该方法能够有效克服目标遮挡、摄像机运动等多种复杂条件的影响,准确并且实时地跟踪运动目标。

## 2 Mean Shift 算法与目标位置概率图

### 2.1 Mean Shift 算法

Mean Shift 是一种典型的非参数模式搜索算法,它能够沿着特征空间的梯度方向进行迭代运算,最终迅速找到数据的局部最优值,在图像分割、聚类、视觉跟踪等领域应用广泛<sup>[19-20]</sup>。运用 Mean Shift 算法进行跟踪主要有两种方式:(1)在候选区域中创建目标位置概率图,利用 Mean Shift 搜索到概率密度最大的位置,如文献<sup>[21]</sup>;(2)通过建立目标相关信息(颜色、梯度等)的分布,在候选区域中通过 Mean Shift 迭代找到与目标分布最相似的区域,如文献<sup>[9]</sup>。第一种方式中的目标位置概率图表示每个像素属于跟踪目标的概率;而第二种方式在迭代时对候选区域的数据进行了加权,其权值也反应了属于跟踪目标的概率,因此两种方式在目标位置概率图上是统一的。但第一种方式的目标位置概率图一旦建立即不会再改变,而第二种方式中的权值则在迭代过程中不断变化。两种方式都面临目标外观变化、遮挡、相似场景干扰、光照突变等问题。

### 2.2 目标位置概率图

目标位置概率图反映了目标在候选区域可能出现的概率分布,如何生成准确的概率图是基于 Mean Shift 的视觉跟踪算法的关键,这些通常以候选区域数据的权值表示目标的位置概率分布。假设初始跟踪目标通过手动方式选择,因此可以

得到以下相关的先验知识:目标区域,目标区域之外的背景区域,以及候选区域 3 种,其中候选区域比目标区域稍大一些。用  $\mathbf{q} = \{q_u \mid u = 1, 2, \dots, m\}$ ,  $\mathbf{p} = \{p_u \mid u = 1, 2, \dots, m\}$ ,  $\mathbf{g} = \{g_u \mid u = 1, 2, \dots, m\}$  分别表示目标、候选区域以及背景区域对应的直方图模型,  $w_i$  表示候选区域像素  $x_i$  对应的权值,  $b(x_i)$  表示像素  $x_i$  对应的直方图区间号,  $\delta[x]$  是克罗内克函数,  $g^*$  表示背景区域直方图分布最小的非零值。文献[15]对当前几种典型的目标位置概率计算方法进行了对比,并提出了对应的复合权值计算方法,表 1 为几种权值计算方法的比较结果。

表 1 权重计算算法对比

Tab.1 Comparison of weight calculation algorithms

Algorithm	Weight
CAMSHIFT[21]	$w_i = \sum_{u=1}^m q_u \delta[b(x_i - u)]$
Mean Shift[9]	$w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u}} \delta[b(x_i - u)]$
Online Selection[10]	$w_i = \sum_{u=1}^m \log \frac{\max(q_u, \epsilon)}{\max(g_u, \epsilon)} \delta[b(x_i - u)]$
Novel Back-Projection[22]	$w_i = \sum_{u=1}^m \frac{\max(q_u, \epsilon)}{\max(g_u, \epsilon)} \delta[b(x_i - u)]$
CBWH[14]	$w_i = \sum_{u=1}^m \sqrt{\min(\frac{g^*}{g_u}, 1)} * \sqrt{\frac{q_u}{p_u}} \delta[b(x_i - u)]$
Fusion-Based[15]	$w_i = \sum_{u=1}^m \frac{\max(q_u, \epsilon)}{\max(g_u, \epsilon)} * \sqrt{\frac{q_u}{p_u}} \delta[b(x_i - u)]$

从表 1 中可以看出,不同的权值计算方法所采用的先验知识并不一致,其准确性与鲁棒性也有明显差异。CAMSHIFT 算法只利用了目标区域信息,因此容易受到与目标具有相似分布的背景的干扰。传统 Mean Shift 采用了目标与候选区域的信息,但没有考虑相似背景区域的干扰。Online Selection 与 Novel Back-Projection 采用了目标与背景区域的信息,然而由于目标与背景区域信息在候选区域容易发生变化,因此这两种算法计算得到的权值准确性较差,容易发生漂移,同时 Novel Back-Projection 没有进行对数运算,其结果容易受到噪声的影响。CBWH 与 Fusion-Based 方法则同时采用了背景、目标以及候选区域的相关信息,区别只在于使用方式不同,且各有优缺点。CBWH 算法所计算的权重对于与目标

区域具有相似分布的背景的抑制并不明显,而 Fusion-Based 算法则容易受到噪声干扰。

由此可知,在不同条件下可以采用不同类型的权值计算方式,以达到最优的组合效果。本文在 Fusion-Based 的基础上提出了一种新的复合权值计算方法,并以此为基础设计了一种多策略融合的视觉目标跟踪算法。

### 3 基于分段复合权值和多策略融合的运动目标跟踪

实现视觉目标跟踪算法是一个多阶段过程,即除了对目标位置的预测之外,还需要对目标的外观模型进行更新以保证目标跟踪的稳定性。本文基于分段复合权值,提出一种多策略融合的运动目标跟踪算法,其基本流程如图 1 所示。

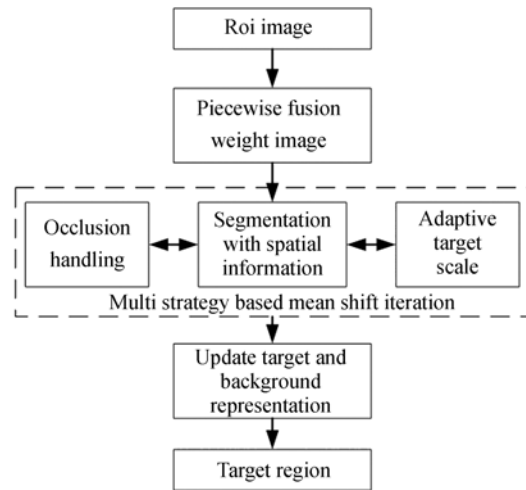


图 1 本文算法的基本流程

Fig.1 Flowchart of proposed algorithm

#### 3.1 分段复合权值

为了充分利用与跟踪目标有关的先验信息,同时保证权值的稳定性与可靠性,本文提出了一种新的分段复合权值计算方法,其具体步骤如下:

(1)当  $p_u < \epsilon_p$  时,表明在候选区域均没有出现对应的特征信息,此时不论目标区域还是背景区域是否有对应特征,权值  $w_u$  均应为 0;若对应的目标信息  $q_u > \epsilon_q$  或者背景  $g_u > \epsilon_g$  则表明跟踪区域出现了偏差,目标处于遮挡或者是消失状态;

(2) 当  $p_u \geq \epsilon_p, q_u < \epsilon_q$  时, 场景中对应新的目标或者是背景区域出现, 但由于不能确认是否为目标, 此时对应的权值  $w_u$  仍然为 0;

(3) 当  $p_u \geq \epsilon_p, q_u \geq \epsilon_q, g_u \geq \epsilon_g$  时, 认为目标、背景、候选区域信息可靠, 按下式计算得到对应的权值:

$$w_u = \sqrt{\min\left(\frac{g_u^*}{g_u}, 1\right)} * \sqrt{\frac{q_u}{p_u}} * \log\left(\frac{q_u}{g_u}\right); \quad (1)$$

(4) 当  $p_u \geq \epsilon_p, q_u \geq \epsilon_q, g_u < \epsilon_g$  时, 认为背景信息并不可靠, 而目标与候选区域信息可靠, 按下式计算得到对应的权值:

$$w_u = \sqrt{\min\left(\frac{g_u^*}{g_u}, 1\right)} * \sqrt{\frac{q_u}{p_u}}, \quad (2)$$

式中  $\epsilon_p, \epsilon_q, \epsilon_g$  分别为分段阈值, 常取为 0.001。通过将  $w_u$  映射到候选区域即可得到对应的目标位置概率分布图  $w_i$ :

$$w_i = \sum_{u=1}^m w_u \delta[b(x_i - u)]. \quad (3)$$

将目标、背景、候选区域按上述方式进行分段, 可以有效减少背景信息以及计算误差的干扰, 并且加快了复合权值的计算速度, 图 2 为 4 种方法对应的权值计算结果, 为了方便显示将权值计算结果均归一化到  $[0, 255]$ 。

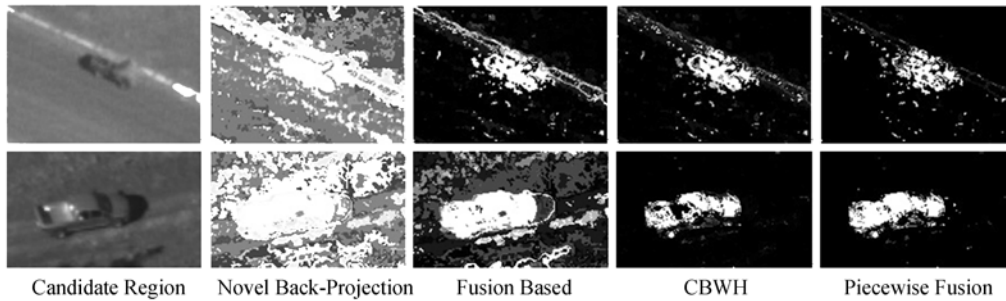


图 2 4 种权值计算方法效果对比

Fig. 2 Comparison of four kinds of weight calculated methods

从图 2 可以看到, CBWH 算法与本文的分段复合权值对计算误差和背景干扰的抑制效果均较好, 可以有效地反映目标的概率分布。而 Novel Back-Projection 与 Fusion Based 方法的结果中则包含了较多的背景信息, 容易导致目标位置的偏移。

### 3.2 多策略融合的 Mean Shift 迭代跟踪

在利用分段复合权值得到关于目标的位置概率图之后, 采用 Mean Shift 进行迭代搜索可以找到目标的中心位置。传统的 Mean Shift 跟踪方法容易受到场景噪声与遮挡的干扰, 并且跟踪区域大小固定, 当目标尺度发生变化时容易造成跟踪漂移, 因此本文融合多种策略以增强 Mean Shift 迭代跟踪的准确性与鲁棒性。

#### 3.2.1 结合空间信息的分割

虽然分段复合权值表示的目标位置概率图已经能较为准确地反映目标位置概率分布, 但为了

进一步抑制场景的干扰, 需要对目标位置概率图进行分割。在将目标位置概率图归一化到  $[0, 255]$  之后, 许多自适应的阈值算法均可应用, 如 Otsu 法、最小交叉熵算法等, 但这些算法并没有考虑到候选区域的空间特性<sup>[23-24]</sup>。本文通过设置滞后阈值并结合空间调整来实现目标位置概率图的自适应分割, 具体步骤如下:

(1) 以  $T_h, T_l$  表示高低阈值,  $S_1, S_2$  表示分割结果, 对目标位置概率图进行初步分割, 如式(4)所示:

$$\begin{cases} S_1 = 255 & w_i \geq T_h \\ S_2 = w_i & w_i < T_h \end{cases}; \quad (4)$$

(2) 利用低阈值对  $S_2$  中像素进行再次分割, 同时进行空间一致性检验, 如式(5):

$$if \begin{cases} w_i \geq T_l \\ \#(R_i \in S_1) > 3 \end{cases} \Rightarrow S_3 = 255, \quad (5)$$

式中:  $R_i$  表示  $w_i$  的邻域像素;  $\#(R_i \in S_1)$  表示  $R_i$  中属于目标区域像素的个数,  $S_3$  表示再次的分割

结果;

(3)将  $S_1$  与  $S_3$  进行组合,得到最终的目标区域  $S_4 = S_1 \cup S_3$ 。

### 3.2.2 遮挡处理

运动目标在跟踪过程中容易受到场景条件的限制,出现局部遮挡或者是完全遮挡的情况,对此跟踪算法必须做出相应的调整,以防止目标丢失。进行跟踪策略调整之前,需要对是否出现遮挡情况进行判定。结合分段复合权值的计算过程,本文通过目标、背景以及候选区域之间的存在关系来判定是否出现遮挡,以  $\eta$ 、 $\theta$  分别表示目标或是背景在与不在候选区域中的直方图区间数,则:

(1)当  $p_u \geq \epsilon_p$  时,目标或背景在候选区域中出现,则  $\eta = \eta + 1$ ;

(2)当  $p_u < \epsilon_p$  时,目标或背景在候选区域中消失,即可能出现遮挡,此时  $\theta = \theta + 1$ ;

(3)若  $\eta < \theta$  则认为出现了明显遮挡,反之则没有。

当跟踪过程中出现遮挡时,迭代跟踪过程仍然将上一帧的位置作为初始位置,但候选区域则变为原来的  $\kappa$  倍。通过扩大候选的方式,可以在目标从遮挡中恢复时,再次捕捉到目标的位置。 $\kappa$  的取值通常可以随着时间递增,直到整个图像区域。

### 3.2.3 尺度自适应

随着目标与摄像机位置的相对变化,目标在图像中的尺度也会相应发生变化,如果对应的感兴趣区域始终保持不变,则容易导致跟踪区域出现漂移,而失去目标。文献[25]利用目标权值图的矩特征提出了一种尺度自适应的 Mean Shift 跟踪算法,但该算法对目标权值的依赖性较大,当候选区域中存在与目标大小相似的背景区域时,其准确度相对较差。文献[26]证明了可以通过目标位置权重图的对数似然函数的下限来决定目标尺度,然而受目标权重图以及固定偏移常量的影响,计算得到的尺度容易发生剧烈变化,并导致跟踪失败。本文在文献[26]的基础上结合历史尺度信息对其进行改进,其具体步骤如下:

(1)以  $h_x^m$ 、 $h_y^m$  分别表示过去  $N$  帧图像中目标最大水平与垂直尺度值,若满足式(6),表明目标

的尺度并没有发生突变,反之则转入步骤(2)。

$$\left\{ \begin{array}{l} \left| h_x^m - \sqrt{\frac{2 \sum_{i=1}^n \omega_i (x_i - c_x)^2}{\sum_{i=1}^n \omega_i}} \right| \leq \varphi \\ \left| h_y^m - \sqrt{\frac{2 \sum_{i=1}^n \omega_i (y_i - c_y)^2}{\sum_{i=1}^n \omega_i}} \right| \leq \varphi \end{array} \right., \quad (6)$$

其中: $n$ 表示跟踪区域的像素总数; $\varphi$ 表示阈值偏差; $c_x$ 、 $c_y$ 表示当前跟踪窗口的中心位置; $x_i$ 、 $y_i$ 表示图像中各像素的位置。目标的水平与垂直尺度由式(7)得到,

$$\left\{ \begin{array}{l} h_x = \frac{1}{2} \left[ \sqrt{\frac{2 \sum_{i=1}^n \omega_i (x_i - c_x)^2}{\sum_{i=1}^n \omega_i}} + \overline{h_x} \right] \\ h_y = \frac{1}{2} \left[ \sqrt{\frac{2 \sum_{i=1}^n \omega_i (y_i - c_y)^2}{\sum_{i=1}^n \omega_i}} + \overline{h_y} \right] \end{array} \right., \quad (7)$$

其中: $\overline{h_x}$ 、 $\overline{h_y}$ 分别表示过去  $N$  帧图像中目标在水平与垂直方向的尺度均值; $h_x$ 、 $h_y$ 表示目标的当前尺度值,同时利用当前值对  $h_x^m$  和  $h_y^m$  进行更新。

(2)当  $h_x^m$  和  $h_y^m$  不满足式(6)时,此时认为目标跟踪结果受到了噪声的干扰,使得计算尺度出现了突变,因而当前尺度由式(8)得到:

$$\left\{ \begin{array}{l} h_x = \frac{\overline{h_x} + h_x^l}{2} \\ h_y = \frac{\overline{h_y} + h_y^l}{2} \end{array} \right., \quad (8)$$

其中: $h_x^l$  和  $h_y^l$  分别表示上一帧目标的水平与垂直尺度。图 3(彩图见期刊电子版)为本文方法与文献中方法的目标尺度自适应估计结果,对比时都采用本文的分段权值作为输入,其中第二行为本文算法结果,第三行为文献的结果,红色框代表目标区域,而绿色框表示搜索区域。

从图 3 可以看出,两种方法都能实现目标尺度的自适应估计,但本文方法对目标尺度的估计更准确与稳定,如第 900 frame 时,原始方法由于尺度偏差累积变化,最终出现了错误的尺度估计结果。

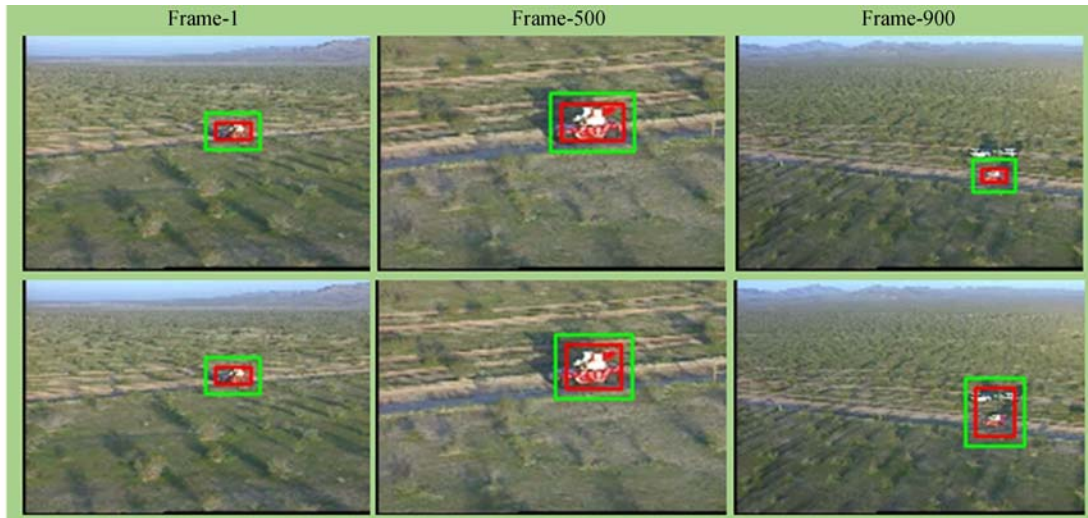


图 3 两种尺度自适应方法结果对比

Fig. 3 Comparison of two different adaptive scale methods

### 3.3 目标与背景信息动态更新

为了应对目标与场景之间的相对变化,需要及时更新目标与背景的特征,比较常用的更新方式是将当前跟踪区域的结果与上一帧的结果直接进行线性加权。然而这种方式对于目标外观的突然变化以及遮挡情况并不加以区别,容易将背景信息更新到目标的直方图中,导致跟踪结果出现偏差。文献[27]采用卡尔曼滤波与假设检测相结合对目标直方图的每个区间进行自适应更新,但并没有讨论背景信息的更新。为了区别目标与背景信息,文献[28]分别对目标与背景信息的更新方式进行了分析,但其信息来源于准确的背景与目标分割。本文结合遮挡情况分析以及特征信息变化对目标与背景直方图进行动态更新,具体的更新过程如下:

(1)以  $q_u^{t+1}$ 、 $q_u^t$ 、 $q_u^{t-1}$  分别表示  $t+1$ 、 $t$ 、 $t-1$  时刻目标直方图,  $g_u^{t+1}$ 、 $g_u^t$ 、 $g_u^{t-1}$  分别表示  $t+1$ 、 $t$ 、 $t-1$  时刻的背景直方图,  $q_u^0$  表示目标的初始直方图,则当没有遮挡情况发生时,目标与背景按式(9)、式(10)进行更新。

$$\begin{cases} q_u^{t+1} = \frac{q_u^{t-1} + q_u^t + q_u^0}{3}, (q_u^t > g_u^t) \wedge (g_u^t = 0) \\ q_u^{t+1} = \frac{q_u^{t-1} + q_u^t}{2}, q_u^t = g_u^t \end{cases}, \quad (9)$$

$$g_u^{t+1} = \frac{g_u^{t-1} + g_u^t}{2}, \quad (10)$$

$$(q_u^t = g_u^t) \vee (g_u^t > q_u^t \wedge q_u^t = 0)$$

(2)当有遮挡情况发生时,只更新背景信息,而不更新目标信息。背景更新时直接将当前场景的背景信息作为下一帧的背景信息,即  $g_u^{t+1} = g_u^t$ 。

## 4 实验结果

### 4.1 评价指标与实验条件

运动目标跟踪算法常用的评测指标包括正确跟踪率(TA)和跟踪框重合率(AOR)<sup>[29]</sup>,当跟踪框重合率大于一定值时,认为跟踪结果正确,因此TA取决于跟踪框重合率。以A表示跟踪结果框,B表示真实的包围框,N表示测试视频总帧数, $M_i$ 表示每一帧的跟踪结果,则TA与AOR计算如式(11)、(12)所示:

$$AOR = \frac{A \cap B}{A \cup B}, \quad (11)$$

$$TA = \frac{\sum_{i=1}^N M_i}{N}, M_i = \begin{cases} 1, AOR \geq 0.5 \\ 0, AOR < 0.5 \end{cases}. \quad (12)$$

实验时选择视觉跟踪常用数据集的4组图像序列对所提算法进行测试验证,分别是 Bolt、FaceOcc1、Jogging 和 Woman<sup>[30]</sup>。这4组视频人工标定了感兴趣目标的跟踪结果,供学者们进行对比分析。表2为实验数据集的具体特性。

表 2 对比实验的 4 段视频序列

Tab. 2 Four Video Sequence for comparative evaluation

Video sequence	Resolution	length	Object size	Characteristic
Bolt	640×360	350	26×61	Occlusion, Deformation, Moving camera
FaceOcc1	352×288	892	114×162	Rotation Occlusion
Jogging	352×288	308	37×114	Occlusion, Deformation, Rotation
Woman	352×288	597	21×95	Occlusion, Illumination Variation, Deformation, Fast motion

本文的运动目标跟踪算法基于 OPENCV 2.4.3 和 Visual Studio2010 环境, 硬件配置为 2.27 GHz 的 Intel Core2Duo CPU 以及 2GB RAM, 上述步骤中的分割阈值  $T_h$  和  $T_l$  分别设置为 200 和 100, 尺度阈值  $\varphi$  设置为 10, 直方图特征采用  $16 \times 16 \times 16$  的 RGB 空间联合直方图。同时选取了文献[14]的 CBWH 算法, 文献[22]的 Novel Back-Projection 算法以及文献[15]的 Fusion Based 算法与本文算法进行对比。

#### 4.2 实验结果与分析

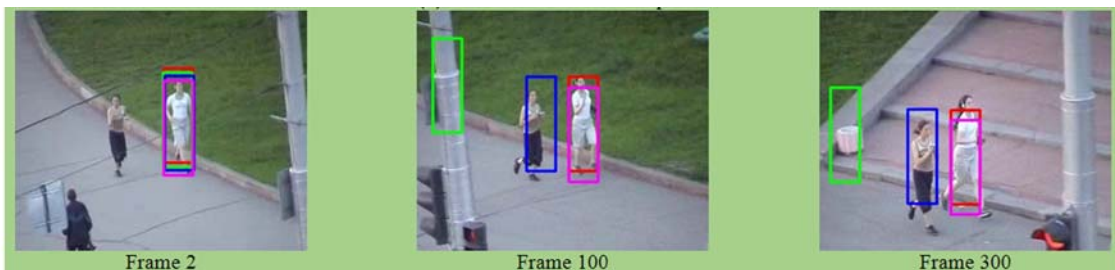
图 4(彩图见期刊电子版)为本文算法与其他几种算法在 4 种场景中的运动目标跟踪结果。从图 4(a)可以看到本文算法对于快速运动目标能够较为准确地进行跟踪, 并且对场景中的相似背景、目标旋转运动、其他相似目标的干扰不敏感; 图 4(b)中虽然目标被书本局部遮挡, 但由于目标



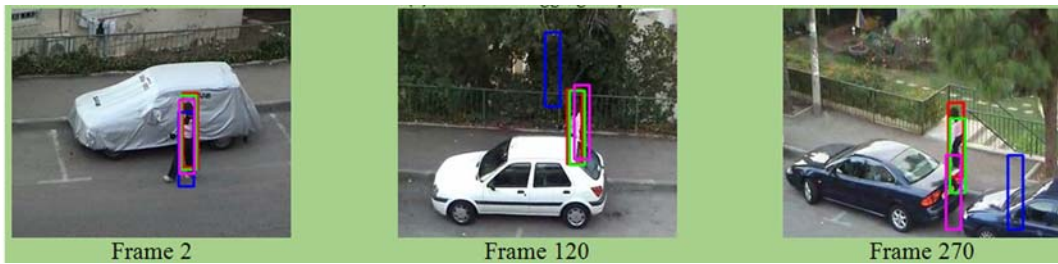
(a) Results of Bolt sequence



(b) Results of FaceOcc1 sequence



(c) Results of Jogging sequence



(d) Results of Woman sequence

图 4 4 种算法的运动目标跟踪结果

Fig. 4 Target tracking results of four algorithms

表 3 4 种算法的性能比较

Tab. 3 Performance comparison of four algorithms

Alogrithm	AOR				TA			
	Bolt	FaceOcc1	Jogging	Woman	Bolt	FaceOcc1	Jogging	Woman
Proposed	0.667	0.736	0.759	0.673	0.791	1.000	0.977	0.894
CBWH	0.409	0.796	0.135	0.666	0.499	1.000	0.167	0.884
Novel Back-Projection	0.170	0.650	0.133	0.088	0.117	0.906	0.170	0.077
Fusion Based	0.306	0.781	0.731	0.444	0.387	1.000	0.944	0.483

与背景的变化相对较小,4 种算法均能较好地跟踪到目标;图 4(c)中目标出现了短暂的完全遮挡现象,采用复合方式的本文算法与 Fusion Based 算法能够较好地跟踪目标,但本文算法的跟踪结果显然更为准确;图 4(d)中目标不仅被局部遮挡同时还受到光照变化以及相似背景的干扰,对此本文算法与 CBWH 算法能够较好地跟踪目标。总体上来看本文算法对目标遮挡具有较强的鲁棒性。

表 3 为 4 种算法的各项评价指标结果,从表中可以看到,除了在 FaceOcc1 视频序列中目标跟踪的重合率之外,本文算法的各项指标均为最优,体现了算法的良好性能。实验时各组视频的平均处理时间约为 10 ms,能够满足实时处理的需求。

#### 参考文献:

- [1] FAN J L, SHEN X H, WU Y. Scribble tracker: a matting-based approach for robust tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(8):1633-1644.
- [2] YANG H, SHAO L, ZHENG F, et al.. Recent

## 5 结 论

本文在分析基于 Mean Shift 框架运动目标跟踪和目标位置概率图计算方法的基础上,提出了一种基于分段复合权值与多策略融合的视觉跟踪算法。该算法首先通过新的分段复合的权值表示目标位置概率图;然后结合目标位置概率图分割、遮挡处理以及目标尺度自适应变化策略进行 Mean Shift 迭代跟踪;最后,利用当前的目标跟踪结果,对目标以及背景区域的特征信息进行更新。通过实验比较,本文算法性能优势明显,对目标遮挡与摄像机运动等因素不敏感,能够在复杂场景下有效地跟踪运动目标并具有较好的实时性。下一步研究将考虑利用场景上下文信息提高目标对相似背景与目标的抗干扰能力。

advances and trends in visual tracking: A review [J]. *Neurocomputing*, 2011, 74(18): 3823-3831.

- [3] COMANICIU D, RAMESH V, MEER P. Real-time tracking of non-rigid objects using mean shift [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Washington DC: IEEE Computer Society, 2000: 142-149.

- [4] 王丽佳,贾松敏,王爽,等.采用改进 Mean Shift 算法的移动机器人行人跟踪[J].光学精密工程,2013,21(9):2364-2370.  
WANG L J, JIA S M, WANG SH, *et al.*. Person tracking of mobile robot using improved Mean Shift [J]. *Opt. Precision Eng.*, 2013, 21(9): 2364-2370.
- [5] WENG S K, KUO C M, TU S K. Video object tracking using adaptive Kalman filter [J]. *Journal of Visual Communication and Image Representation*, 2006, 17(6): 1190-1208.
- [6] ISARD M, BLAKE A. Condensation-conditional density propagation for visual tracking [J]. *International Journal of Computer Vision*, 1998, 29(1):5-28.
- [7] 宋策,张葆,尹传历.适于机载环境对地目标跟踪的粒子滤波设计[J].光学精密工程,2014,22(4):1037-1047.  
SONG C, ZHANG B, YIN CH L. Particle filter design for tracking ground targets in airborne environment [J]. *Opt. Precision Eng.*, 2014, 22(4): 1037-1047.
- [8] 龚俊亮,何昕,魏仲慧,等.采用改进辅助粒子滤波的红外多目标跟踪[J].光学精密工程,2012,20(2):413-421.  
GONG J L, HE X, WEI ZH H, *et al.*. Multiple infrared target tracking using improved auxiliary particle filter [J]. *Opt. Precision Eng.*, 2012, 20(2):413-421.
- [9] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5):564-577.
- [10] COLLINS R, LIU Y, LEORDEANU M. Online selection of discriminative tracking features [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1631-1643.
- [11] BIRCHFIELD S T, RANGARAJAN S. Spatio-grams versus histograms for region-based tracking [C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington DC: IEEE Computer Society, 2005: 1158-1163.
- [12] JEYAKAR J, BABU R V, RAMAKRISHNAN K R. Robust object tracking with background-weighted local kernels [J]. *Computer Vision and Image Understanding*, 2008, 112(3): 296-309.
- [13] LI S X, CHANG H X, ZHU C F. Adaptive pyramid mean shift for global real-time visual tracking [J]. *Image and Vision Computing*, 2010, 28(3): 424-437.
- [14] NING J, ZHANG L, ZHANG D, *et al.*. Robust mean-shift tracking with corrected background-weighted histogram [J]. *IET Computer Vision*, 2012, 6(1): 62-69.
- [15] WANG L, PAN C, XIANG S. Mean-shift tracking algorithm with weight fusion strategy [C]. *Proceedings of 18th IEEE International Conference on Image Processing*, New York, USA: IEEE Press, 2011: 473-476.
- [16] CHOI H S, KIM I S, CHOI J Y. Combining histogram-wise and pixel-wise matchings for kernel tracking through constrained optimization [J]. *Computer Vision and Image Understanding*, 2014, 118: 61-70.
- [17] 刘伟宁.基于 DSP+FPGA 平台的复杂背景目标检测与跟踪[J].液晶与显示,2014,29(6):1151-1155.  
LIU W N. Object tracking under complicated background based on DSP+FPGA platform [J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(6):1151-1155.
- [18] 贺柏根,刘剑,刘伟宁,等.基于 TMS320C6455+FPGA+SDRAM 的快速视频跟踪系统设计[J].液晶与显示,2014,29(6):1111-1116.  
HE B G, LIU J, LIU W N, *et al.*. Fast image tracking system design based on TMS320C6455+FPGA+SDRAM[J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(6):1111-1116.
- [19] CHENG Y. Mean shift, mode seeking, and clustering [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, 17(8): 790-799.
- [20] COMANICIU D, MEER P. Mean shift: A robust approach toward feature space analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 603-619.
- [21] BRADSKI G. Computer vision face tracking for use in a perceptual user interface [C]. *Proceedings of IEEE Workshop On Applications of Computer Vision*, New Jerseys, USA: IEEE Press, 1998: 214-219.
- [22] WANG L F, WU H Y, PAN C H. Mean-Shift object tracking with a novel back-projection calcu-

- lation method[C]. *Proceedings of the 9th Asian Conference on Computer Vision-Volume, Berlin, Germany, Springer-Verlag Press*, 2009: 83-92.
- [23] OTSU N. A threshold selection method from gray-level histogram [J]. *IEEE Trans. on Systems, Man, and Cybernetics*, 1979, 9: 62-66.
- [24] BRINK A D, PENDOCK N E. Minimum cross-entropy threshold selection [J]. *Pattern Recognition*, 1996, 29(1): 179-188.
- [25] NING J, ZHANG L, ZHANG D, *et al.*. Scale and orientation adaptive mean shift tracking [J]. *IET Computer Vision*, 2012, 6(1): 52-61.
- [26] JIANG Z L, LI S F, GAO D F. An adaptive mean shift tracking method using multiscale images [C]. *Proceedings of International Conference on Wavelet Analysis and Pattern Recognition*, New Jerseys, USA: IEEE, 2007: 1060-1066.
- [27] PENG N S, YANG J, LIU Z. Mean shift blob tracking with kernel histogram filtering and hypothesis testing[J]. *Pattern Recognition Letters*, 2005, 26(5): 605-614.
- [28] FAN J, SHEN X, WU Y. Closed-loop adaptation for robust tracking [C]. *Proceedings of the 11th European Conference on Computer Vision, Berlin, Germany, Springer-Verlag Press*, 2010: 411-424.
- [29] SEVILLA-LARA L, LEARNED-MILLER E. Distribution fields for tracking [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, New York, USA: IEEE Press, 2012: 1910-1917.
- [30] WU Y, LIM J, YANG M H. Online object tracking: A benchmark [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, New York, USA: IEEE Press, 2013: 2411-2418.

#### 作者简介:



苏延召(1987—),男,四川资阳人,博士研究生,2008年、2010年于第二炮兵工程学院分别获得学士、硕士学位,主要从事机器视觉及自动控制方面的研究。  
E-mail: syzlh@163.com

#### 导师简介:



李艾华(1966—),男,河北沧州人,博士,教授,博士生导师,1987年、1990年于第二炮兵工程学院分别获得学士、硕士学位,1998年于西安交通大学获得博士学位,现为第二炮兵工程大学阵地管理工程教研室主任,主要从事机电设备故障诊断、自动控制、机器视觉等方面的研究。E-mail: ryulong@126.com

(版权所有 未经许可 不得转载)