

文章编号: 1672-2892(2010)06-0673-05

含区域信息的核函数物体跟踪算法

孙 研, 李才会, 陆起涌

(复旦大学 电子工程系, 上海 200433)

摘要:“无穷多解”问题, 是基于核函数的视频物体跟踪算法的主要问题。而“多极值点”问题相比“无穷多解”而言更为普遍, 却尚未被关注。本文对上述问题进行了研究, 并且提出了一种含区域信息的核函数物体跟踪算法——区域化核函数跟踪算法(SBT)。该方法主要利用物体分割获得的区域权重信息来消除虚假的极值点, 从而使跟踪过程有更高的鲁棒性。此外, 该新方法自身天然具有处理物体缩放及旋转的能力。实验证明, SBT 算法在跟踪过程中表现出了更好的性能。

关键词: 核函数; 物体跟踪; 区域信息; 误差平方和

中图分类号: TN820.4

文献标识码: A

Kernel-based tracking with section information

SUN Yan, LI Cai-hui, LU Qi-yong

(Electronic Engineering Department, Fudan University, Shanghai 200433, China)

Abstract: The well-known “singularity” problem is one of the major defects in kernel-based video object tracking. However, the multiple-extreme problem is more general but has not been paid attention to. This paper studies this important problem and presents a novel approach called Section-Based Tracking (SBT) that is a kernel-based method with the section information provided by the division of the object’s weight image. SBT serves to eliminate fake extreme points and make the tracking process more robust. Besides, this new approach can also be naturally extended to deal with scale change and rotation. Experiments show that SBT method conduces to better performance in tracking process.

Key words: kernel; object tracking; section information; Sum of Squared Differences(SSD)

基于核函数的视频物体跟踪算法^[1]由于其出色的性能尤其是较低的运算复杂度近年来吸引了大家的广泛注意。经典的核函数跟踪算法用 Epanechnikov 核将物体描述为按色彩分类的各像素点权重和的直方图, 每个像素点的权重大小由它自身在图像平面的位置决定。此外, 该算法采用基于 Bhattacharyya 系数的测度^[1]来评价模板和实际目标之间的差异, 并通过均值漂移(mean-shift)^[2-3]的方法最小化差异来实现对物体的跟踪。

核函数视频物体跟踪算法的主要问题之一是大家熟知的“无穷多解”问题, 即通过观察值无法唯一确定物体的当前位置。George Hager 通过引入误差平方和(SSD)测度和牛顿迭代详细阐明了核函数算法的这个缺陷^[4]。Hager 将这个缺陷的原因归结为在牛顿迭代过程中方程的系数矩阵的秩不足, 并且提出了用多核的方法来解决这个问题。Fan 进一步使这些多核之间相互限制, 相互配合^[5], 而 Wei Qu 通过利用系统的可观测性理论增强了核函数的鲁棒性^[6]。核函数跟踪的其他相关问题, 比如物体尺度变化、旋转、背景干扰等也引起了很多人的兴趣。Collins 提出了尺度核来处理物体大小的变化^[7], Yilmaz 使用基于水平集^[8]的非对称核来同时处理物体的缩放和旋转^[9], 而陈向东等通过改进的直方图映射提高了抗背景干扰的能力^[10]。除了以上关注点, 在核函数物体跟踪过程中还存在一个经常出现却没有被提出过的问题——多极值点。在某些情况下, 尽管“无穷多解”的情况没有出现, 目标的位置可以被唯一确定, 但是得到的这个位置可能是错误和虚假的, 这是因为存在不止一个局部极值点, 干扰了算法的判断。为了解决这个问题, 笔者在 Hager 提出的 SSD 算法的基础上, 充分利用目标的权重分布信息, 提出了一个新方法: 区域化核函数跟踪算法(SBT)。分析和实验都表明, 即使在单核情况下, SBT 在应对“无穷多解”以及更为一般的“多极值点”问题时, 都表现出色, 且自身天然具备处理目标缩放及旋转的能力。此外, 笔者设计了一个新核“QPeak”, 使得包括 SBT 在内的核函数跟踪过程更具鲁棒性。

1 SSD 跟踪算法

SSD 核函数跟踪算法^[4]的性能比广为人知的均值漂移核函数跟踪算法更为出色。与经典的核函数跟踪算法一致,物体被描述为按色彩分类的各像素点权重和的直方图,而权重的大小由像素点在图像平面内的位置决定:

$$q_u = \frac{1}{C} \sum_{i=1}^n K(\mathbf{x}_i - \mathbf{c}) \delta(b(\mathbf{x}_i), u) \quad (1)$$

式中: q_u 表示按色彩分类的权重和直方图上的一个类别(即直方); $\{\mathbf{x}_i, i=1, 2, \dots, n\}$ 是像素点在图像平面的位置; n 为像素点个数; $b(\mathbf{x}_i)$ 是分类函数,分类标准可以是色彩信息或者灰度信息等; u 为特征数值; δ 是克罗内克函数(Kronecker delta); K 是如 Epanechnikov 等的核函数,以物体内某一点 \mathbf{c} 为中心(通常是物体中心点),用于计算像素点的权重; C 为归一化项,使得 $\sum_{u=1}^m q_u = 1$ 。用矩阵形式完整表达,得到:

$$\mathbf{q} = \mathbf{U}^T \mathbf{K}(\mathbf{c}) \quad (2)$$

其中, t 为时间标号,而

$$\mathbf{U} = \begin{pmatrix} \delta(b(\mathbf{x}_1), u_1) & \dots & \delta(b(\mathbf{x}_1), u_m) \\ \vdots & \ddots & \vdots \\ \delta(b(\mathbf{x}_n), u_1) & \dots & \delta(b(\mathbf{x}_n), u_m) \end{pmatrix} \in \mathbf{R}^{n \times m} \quad (3)$$

$$\mathbf{K}(\mathbf{c}) = \frac{1}{C} (K(\mathbf{x}_1 - \mathbf{c}) \cdots K(\mathbf{x}_n - \mathbf{c}))^t \in \mathbf{R}^n \quad (4)$$

由此,物体模板和候选区域可分别写为 $\mathbf{q} = \mathbf{U}^T \mathbf{K}(\mathbf{c})$ 和 $\mathbf{p}(\mathbf{c}) = \mathbf{U}^T \mathbf{K}(\mathbf{c}')$ 。跟踪问题就归结为找到核函数的中心点在新一帧中的位置 \mathbf{c}' ,使得物体模板和候选区域间的差异最小。在 SSD 核函数跟踪算法中,该差异用误差平方和(SSD)来定义:

$$O(\mathbf{c}) = \|\sqrt{\mathbf{q}} - \sqrt{\mathbf{p}(\mathbf{c})}\|^2 \quad (5)$$

与经典的均值漂移相比,SSD 跟踪算法中所采用的牛顿迭代是一种更为高效的方法,有着相当出色的一步收敛性能。在 SSD 测度下,牛顿迭代的一步解的方程如下:

$$\mathbf{J}_K^t \mathbf{U} \mathbf{d}(\mathbf{p}(\mathbf{c}))^{-1} \mathbf{U}^T \mathbf{J}_K \Delta \mathbf{c} = 2 \mathbf{J}_K^t \mathbf{U} \mathbf{d}(\mathbf{p}(\mathbf{c}))^{-1/2} (\sqrt{\mathbf{q}} - \sqrt{\mathbf{p}(\mathbf{c})}) \quad (6)$$

$$\mathbf{J}_K = (\partial \mathbf{K} / \partial c_1, \partial \mathbf{K} / \partial c_2) \quad (7)$$

(c_1, c_2) 为向量 \mathbf{c} 即核函数中心点在图像平面的坐标,而 $\mathbf{d}(\mathbf{p})$ 表示以 \mathbf{p} 为对角线的对角阵。

SSD 算法使得跟踪过程更为快速,大大提高了核函数跟踪算法的性能,然而像多极值点问题等仍然会影响跟踪的准确性,这使得笔者尝试进一步提高算法的稳定性。

2 多极值点问题及区域化核函数跟踪算法

2.1 多极值点问题

与由“秩不足”造成的“无穷多解”相比,多极值点问题在核函数跟踪算法中更为普遍却未被提出。实际上,无穷多解现象只是多极值点问题的一种。为了便于讨论且不失一般性,接下来的论述将基于 Hager 提出的性能优异的 SSD 核函数跟踪算法。所谓多极值点,是指物体模板和候选图像区域间的差异,如误差平方和(SSD)存在不止一个局部极值点,由此导致牛顿迭代或其他收敛算法的结果不准确或者不唯一。多极值点的情况经常有可能发生,尤其是在多个物体间比较相像或者物体有一部分与背景类似的时候。

一个简单的例子如图 1 所示。使用单个 Epanechnikov 核来计算目标的权重直方图。在图 1(a)中的矩形框表示需要跟踪的目标, c_1 是实际选择的核函数中心或者说是 SSD 的真实极值点。然而,由于黑色的背景与物体的黑色部分颜色一致,在图像平面上还存在另一个虚假的极值点 c_2 。如果将核函数中心点移至 c_2 ,即将标示物体的矩形框移至以 c_2 为中心,如图 1 (b)所示,那么将会得到和在 c_1 处完全一样的权重直方图。换句话说, c_1 与 c_2 是两个 SSD 极值相等(均为 0)的极值点。又由于 c_1, c_2 两个点很接近,跟踪

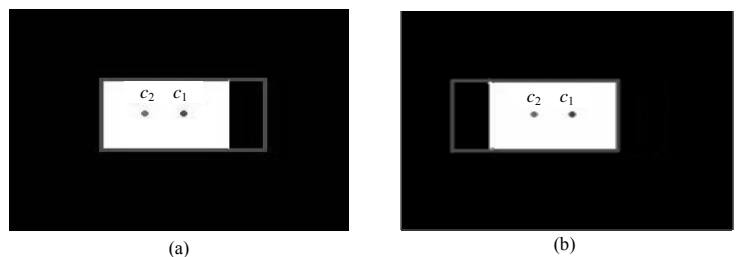


Fig.1 Example of background disturbing
图 1 背景干扰的例子

过程中很容易发生从一个点偏到另一个点的情况, 跟踪变得很不可靠。这个例子不能用秩不足造成的“无穷多解”来解释, 因为核函数的中心仍旧可以用式(6)计算得到唯一解, 尽管得到的值可能是虚假极值点。实际上, “无穷多解”问题是极值点连续分布的多极值点问题, 而本例是极值点离散分布的情况。

2.2 区域化核函数跟踪算法

如 2.1 所述, 多极值点问题会严重干扰核函数物体跟踪过程, 造成跟踪失败。然而, 重新观察一下图 1 会发现, 尽管标示物体的矩形框在以 c_1 为中心和以 c_2 为中心时的权重直方图完全一致, 但是黑色区域与白色区域的相对位置发生了改变, 这使笔者意识到可以在原有 SSD 算法的基础上加入权重的区域分布信息来解决多极值点问题。这里的“区域”, 是指物体在图像平面内的一部分或一个划分。

用 $f_v(x_i)$ 表示区域分割函数, 则式(1)可改写为:

$$q_{uv} = \frac{1}{C} \sum_{i=1}^n K(x_i - c) \delta(b(x_i), u) f_v(x_i) \tag{8}$$

式中:

$$f_v(x_i) = \prod_j \delta[a_j(x_i), v_j] \tag{9}$$

$a_j(x_i)$ 是区域分类函数, 将像素 x_i 所属的区域映射到对应的 v_k 值上。 \mathbf{v} 是一个向量, 表示由一个或多个区域分类函数划分出来的一个特定区域。换句话说, 函数 $a_j(x_i)$ 反映了对目标的一种划分, 而克罗内克函数的积则反映了多种划分的组合。

有了式(8)和式(9), 目标的权重直方图便可以反映像素点所在的区域信息, 对目标的描述也变得更为精确。也可以将式(8)写成矩阵形式:

$$\mathbf{q}_s = \mathbf{U}_s' \mathbf{K}(c) \tag{10}$$

式中:

$$\mathbf{U}_s = [\mathbf{u}_{11}, \mathbf{u}_{21}, \dots, \mathbf{u}_{m1}, \mathbf{u}_{12}, \mathbf{u}_{22}, \dots, \mathbf{u}_{md}] \tag{11}$$

$$\mathbf{u}_{ik} = [\delta(b(x_1), i) f_k(x_1), \dots, \delta(b(x_n), i) f_k(x_n)]' \tag{12}$$

在 \mathbf{u}_{md} 中的下标 d 表示区域分类函数及其组合划分的区域总数; m 与在 SSD 算法中的意义一致, 表示按色彩分类的权重直方图上直方的个数; $f_k(x_i)$ 表示按一定次序排列的一系列 $f_v(x_i)$ 中的一个; \mathbf{q}_s 是描述目标的向量, 向量中的元素 q_{uv} 与 \mathbf{U}_s' 中元素的排列次序相对应。显然, 式(10)与式(2)具有相同的形式, 牛顿迭代过程可以和原先在 SSD 算法中一样地在 SBT 算法中运用。由于 \mathbf{q}_s 的维度大于 \mathbf{q} , 所以有:

$$\text{rank}(d(\mathbf{p}_s)^{-1/2} \mathbf{U}_s' \mathbf{J}_k) \geq \text{rank}(d(\mathbf{p})^{-1/2} \mathbf{U}' \mathbf{J}_k) \tag{13}$$

这表明区域信息有助于增加方程系数矩阵的秩, 从而降低了发生“无穷多解”问题的可能性。同时, 由于区域信息的加入, 更为常见的“多极值点”问题发生的可能性也大为降低。重新考虑一下图 1 的情况, 如果将物体分成左右两个区域, 则由于物体黑色部分所在的区域不同, 此时跟踪过程将不再受虚假极值点 c_2 的干扰。

理论上, 物体区域的划分数量是任意的。然而, 尽管更多的区域能够提供对物体更为精确的描述, 区域数量过多却会导致吸引盆过小, 造成跟踪能力降低。在实际应用中, 建议每个区域的像素点数量平均在 150 以上。

3 SBT 算法扩展

有了区域信息的帮助, 物体的尺度变化和旋转便可以在单个核函数的条件下, 通过区域的运动情况进行计算。按照区域重写式(11), $\mathbf{U}_s = [\mathbf{u}_{s1}, \mathbf{u}_{s2}, \dots, \mathbf{u}_{sd}]$, 其中, $\mathbf{u}_{si} = [u_{1i}, u_{2i}, \dots, u_{mi}]$, 每个 \mathbf{u}_{si} 唯一对应一个区域。同样的, 有相应的 \mathbf{q}_{si} 和 \mathbf{p}_{si} 。用 $\mathbf{q}_{si}, \mathbf{p}_{si}, \mathbf{u}_{si}$ 代替 $\mathbf{q}, \mathbf{p}, \mathbf{u}$ 代入式(6)进行计算, 得到了每个区域的平移参数。在实际应用中, 通过将所有区域对应的 $\mathbf{q}_{si}, \mathbf{p}_{si}, \mathbf{u}_{si}$ 分别按相同的顺序排列, 可以利用式(6)一次计算出所有区域的平移参数。这些区域平移参数同时反映了整个物体的平移以及各区域对于核函数中心点的相对运动。

用 Δc 表示核函数中心的移动, 也就是整个物体的平移参数, 用 Δc_{si} 表示某一个区域的平移参数, 则每个区域对于中心点的相对运动即为 $\Delta c_{ri} = \Delta c_{si} - \Delta c$ 。 Δc_{ri} 反映了物体的缩放及旋转, 在每个区域中定义一个参照点, 将该点的坐标作为该区域的位置, 由此, 可以利用 Δc_{ri} , 在以核函数中心为原点的极坐标下计算各区域相对于原点的距离及角度的变化。距离的变化对应着物体的缩放, 而角度的变化则对应着物体的旋转。

此外,受多个核函数之间约束条件的启发^[5],也为各个区域之间的关系增加了约束条件,以此来进一步提高 SBT 算法的性能。由于各个区域互相紧密连接共同构成整个物体,因此,利用区域间的结构关系添加约束条件是一个比较自然的做法。在实际应用中使用了这样的约束条件,即所有区域对于核函数中心的相对运动的矢量和为零。当然,也可以采用其它合理的约束条件。

4 实验及结果

为了进一步提高核函数跟踪的效果,笔者设计了一个新核——“QPeak”,其表达式如下:

$$K = \lambda(1 - (|x - c_1|/x_{scale}) - 0.5) - (|y - c_2|/y_{scale}) - 0.5 \quad (14)$$

式中: (x, y) 是物体像素点坐标; (c_1, c_2) 是核函数中心点坐标; x_{scale} 和 y_{scale} 分别是物体边缘到核函数中心点的最大水平与竖直距离; λ 是常数。该核函数的好处是显而易见的。

首先,该核函数具有分段线性,因此更适用于牛顿迭代过程,具有更好的一步收敛性能。其二,设 p_u, q_u 分别是目标模板和候选图像的权重直方图的第 u 项,即第 u 个直方,根据式(5),在 p_u, q_u 间的差异一定的情况下,较大的 p_u, q_u 会导致较小的 ΔO_u 。换句话说,SSD 测度对权重和较小的点集更为敏感。与 Epanechnikov 核函数相比, QPeak 有着更多的小权重区域,因此能够对物体的更多部分敏感,进而使得跟踪过程更具鲁棒性。该新核将在接下来的实验中使用。

4.1 物体位置跟踪

在这一部分,将比较 SBT 算法和 SSD 算法在跟踪物体平移过程中的表现。由于 SSD 算法被认为优于均值漂移算法,所以没有加入均值漂移算法作为 SBT 的比较对象。在实验中, SBT 算法分别在采用 Epanechnikov 核(E-SBT), QPeak 核(Q-SBT)以及在采用 QPeak 核的条件下进一步添加约束条件(QC-SBT)这 3 种情况下进行计算; SSD 算法采用 Epanechnikov 核(E-SSD)进行计算。尽管 QPeak 核同样能够提升 SSD 算法的性能,但由于篇幅限制,在本文中没有列出,具体可参见文献[11]。跟踪过程都基于灰度域,权重直方图按彩色转灰度后的灰度值分成 32 类(即直方的个数)。SBT 算法中区域的个数为 4。为了使比较更有说服力,没有使用任何高级的模板更新算法,下一帧图像直接使用当前帧的跟踪结果作为目标模板。每一帧图像处理时的牛顿迭代次数均为 2。

下面列出了实验中具有代表性的 2 个例子。图 2 是对人脸的跟踪结果,人脸与身后作为大块背景的门在灰度域上非常接近,由于算法在灰度域进行,此种情况即为前文讨论的背景干扰的情况。图 3 是对运动员头盔的跟踪,头盔的主体颜色与球衣的部分颜色一致,这是物体间类似的情况。上述 2 种情况均很容易出现多极值点问题,跟踪很容易被灰度相似的背景或物体影响。然而,从图 2 和图 3 反映的实验结果可以看出,带有区域信息、对物体描述更精确的 SBT 算法,表现出了相比 SSD 更为优异的性能,且在 QPeak 核的帮助下,跟踪过程更加准确。而在进一步加入约束条件后,由于 Q-SBT 算法表现已经很不错,没有更多的提升空间,因此 QC-SBT 与 Q-SBT 的实验结果看不出差别,为节省篇幅,用一张图同时表示 Q-SBT 与 QC-SBT 的结果。

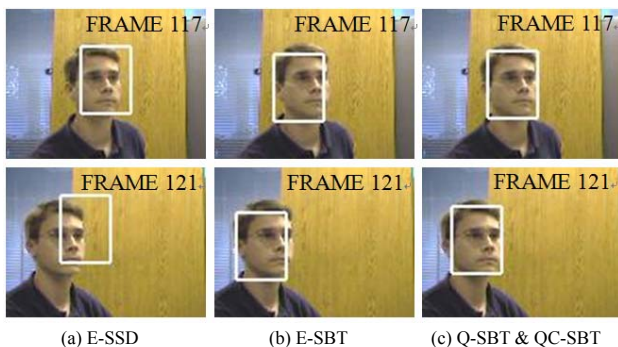


Fig.2 Tracking of human face.
图 2 人脸跟踪

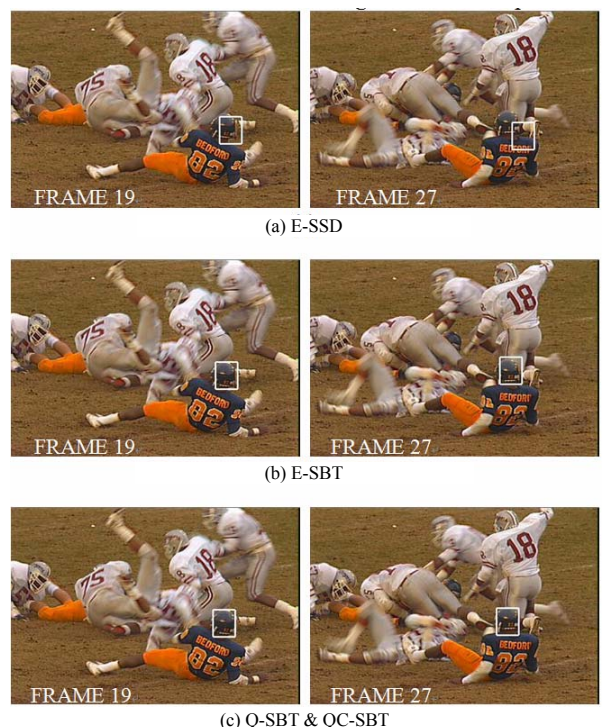


Fig.3 Tracking of the blue head
图 3 蓝色头部跟踪

4.2 物体缩放及旋转

在这一部分给出了物体尺度变化及旋转的跟踪结果,分别是大小发生变化的人脸和一张绕顶点旋转的 CD 盒。计算缩放及旋转的方法即如第 3 部分“SBT 算法扩展”中所述。由于计算物体的旋转及缩放对各区域的平移参数精确度要求较高,因此采用含约束条件的基于 QPeak 核的 SBT 算法,即 QC-SBT 来计算。采用的约束条件为前文所述的结构约束,即所有区域对于核函数中心的相对运动的矢量和为零。物体的区域数和平移跟踪实验一样是 4 个。实验结果如图 4、图 5 所示。很明显,SBT 算法具有很好的处理物体缩放及旋转的能力。



Fig.4 Tracking of human face with scale change
图 4 跟踪大小变化的人脸



Fig.5 Tracking of the rotating CD
图 5 跟踪旋转的 CD

5 结论

本文提出并研究了多极值点问题。为解决这一问题,本文在 SSD 算法的基础上提出了区域化核函数跟踪算法 SBT 及其扩展。实验证明,该算法能够有效减少多极值点问题的发生,增加跟踪过程的鲁棒性,并天然具有处理物体尺度变化及旋转的能力。

参考文献:

- [1] Comaniciu D,Ramesh V,Meer P. Kernel-Based Object Tracking[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2003,25(5):564-575.
- [2] Comaniciu D,Meer P. Mean Shift:A Robust Approach toward Feature Space Analysis[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2002,24(5):603-619.
- [3] Cheng Y. Mean Shift, Mode Seeking, and Clustering[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1995, 17(8):790-799.
- [4] Hager G D,Dewan M,Stewart C V. Multiple Kernel Tracking with SSD[C]// Proc. IEEE Conf. Computer Vision and Pattern Recognition. Washington, DC, USA:[s.n.], 2004:790-797.
- [5] Fan Z,Yang M, Wu Y. Multiple Collaborative Kernel Tracking[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2007,29(7):1268-1273.
- [6] Qu W,Schonfeld D. Robust Control-Based Object Tracking[J]. IEEE Trans. Image Processing, 2008,17(9):1721-1726.
- [7] Collins R T. Mean-Shift Blob Tracking through Scale Space[C]// Proc. IEEE Conf. Computer Vision and Pattern Recognition. Madison, Wisconsin:[s.n.], 2003:234-240.
- [8] Sethian J. Level set methods:evolving interfaces in geometry, fluid mechanics computer vision and material sciences[M]. New York:Cambridge University Press, 1999.
- [9] Yilmaz A. Object Tracking by Asymmetric Kernel Mean Shift with Automatic Scale and Orientation Selection[C]// Proc. IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN:[s.n.], 2007:1-6.
- [10] 陈向东,安国成,吴镇扬. 增强抗背景干扰能力的均值移动目标跟踪算法[J]. 信息与电子工程, 2008,6(1):40-45. (CHEN Xiang-dong,AN Guo-cheng,WU Zhen-yang. Algorithm for Object Tracking Using Improved Histogram Back-projection and Mean-shift with Enhanced Ability of Anti-jamming[J]. Information And Electronic Engineering, 2008,6(1):40-45.)
- [11] Sun Y,Chen X,Lu Q. Robust Kernel Design for SSD Tracker[C]// IEEE Youth Conference on Information, Computing and Telecommunication, 2009. YC-ICT '09. Beijing:[s.n.], 2009:162-165.

作者简介:



孙 研(1985-),男,江苏省无锡市人,在读硕士研究生,主要研究方向为视频跟踪算法。
email:sunyanRobin@gmail.com.

李才会(1986-),女,贵州遵义人,在读硕士研究生,主要研究方向为视频处理。

陆起涌(1966-),男,江苏太仓人,博士,副教授,主要研究方向为视频、嵌入式系统、计算机网络等。