

基于 QP-TR 信任域的序列图像目标跟踪^{*}

贾静平 柴艳妹 赵荣椿 张艳宁

(西北工业大学计算机学院 西安 710071)²

摘要 本文将信任域算法和尺度空间理论相结合,提出了一种能够精确描述目标尺寸连续变化的新的序列图像目标跟踪算法;将信任域算法与灰度模板相结合,提出了一种新的实时目标跟踪算法。在第一种算法中,首先将序列图像按照颜色直方图转换成目标概率分布图,目标区域在概率分布图中呈现为灰度块。然后通过检测该图在尺度空间中微分滤波器输出的极值,来决定这些灰度块的尺度。最后我们使用 QP-TR 信任域算法在尺度空间和图像平面内快速搜索概率分布图的多尺度规范化 Laplacian 滤波函数极值,实现了目标定位并同时决定了其尺度,从而完成了跟踪任务。在第二种算法中,首先记录目标初始模板,在随后每一帧中应用 QP-TR 信任域算法搜索与该模板最相似的区域,实现目标定位。和现有算法的比较以及在大量真实序列图像上的实验表明,两种算法分别在目标大小描述,跟踪精度上以及运算速度上有了显著提高。

关键词 QP-TR 信任域算法,尺度空间,多尺度规范化 Laplacian 滤波,实时目标跟踪

Tracing of Objects in Image Sequences Using QP-TR Trust Region Algorithm

JIA Jing-Ping CHAI Yan-Mei ZHAO Rong-Chun ZHANG Yan-Ning

(School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072)

Abstract A new tracking framework based on the QP-TR trust region algorithm is proposed, in which two independent algorithms appropriate for different situations are demonstrated. In the first algorithm the constant changes of the target's size can be precisely described. For each incoming frame, a probability distribution image of the target is created, where the target's area turns into a blob. The scale of this blob can be determined based on local maxima of differential scale-space filters. We employ the QP-TR trust region algorithm to search the local maxima of multi-scale normalized Laplacian filter of the probability distribution image to locate the target as well as determine its scale. In the second algorithm, we combine the template matching with the QP-TR method and achieved the real time performance. In the presented tracking examples, the two algorithms demonstrate their great improvement on tracking precision and runtime performance respectively.

Keywords QP-TR algorithm, Scale space, Multi-scale normalized laplacian filter, Real-time tracking of objects

序列图像中的目标跟踪算法是视频监控、智能用户接口、基于对象的视频压缩、巡航导弹末端制导和辅助驾驶等许多计算机视觉应用中的关键技术。一个典型的跟踪算法主要包含两个部分:目标特征表示及目标定位。常用的目标特征有灰度值、形状信息、颜色信息、纹理特征以及其速度、加速度等;目标定位则是在各帧中寻找与目标特征最匹配的区域,因而可将其转化为一个循环优化问题。

与当前流行的基于线性搜索的 Mean Shift 方法^[1]相比,信任域算法是一种性能更优良的优化算法。而 QP-TR 算法在经典的信任域算法基础上又进一步改良了获得目标函数的海森矩阵和梯度的方式,因此具有更高的精度和更好的迭代性能^[2]。

跟踪过程中能否准确决定目标的大小直接影响着跟踪的精度。当前已有的算法^[1,3,4]大多是靠试探的方式来决定目标尺寸的,但是这种在尺度空间里取离散值的方式并不能很准确地描述目标的大小。针对这一问题,本文引入了 Lindeberg 提出的尺度空间理论^[5],通过检测尺度空间中微分滤波器输出的极值来自适应地决定特征尺度,在连续的尺度空间里精确刻画目标大小的连续变化。利用 QP-TR 算法在尺度空间里和图像平面内快速搜索多尺度、规范化的 Laplacian 滤波函数的极值来进行目标跟踪,避免了尺度空间的显式产生,从而有效解决了目标定位和尺度决定的问题。

在很多跟踪应用中,实时性是很重要的要求。通常搜索

窗大小和计算量是一对矛盾,增大搜索窗通常也意味着计算量的加大。我们将 QP-TR 算法与灰度模板相结合,成功地扩大了搜索窗,同时也降低了计算量。

1 QP-TR 信任域算法

循环优化算法大致可划分为两类:线性搜索和信任域^[6]。在线性搜索方法中,每次循环都沿着梯度下降最快的方向进行。而信任域方法每次循环是在一个有界区域内搜索一个最优值(在这个有界区域内求最优值的过程称为信任域的子问题)。因此,线性搜索可以看成是信任域方法的一个特例。

1.1 经典信任域算法

经典信任域算法可用来解决下面的无约束最小化问题: $\min_{x \in V} f(x)$ 。其中 f 是向量空间 V 中的目标函数。信任域算法包含三个主要部分:1)信任域半径:决定了信任域的大小;2)信任域子问题:在一个邻域内近似并求目标函数最小值;3)信任域保真度:衡量子问题近似的精确程度。

给定初始点 $x_0 \in R^n$, 初始信任域半径 Δ_0 , 经典信任域算法步骤如下:

①初始化:设定常数 $\eta, \rho, \gamma_1, \gamma_2$, 满足 $0 < \eta \leq \rho < 1, 0 < \gamma_1 \leq \gamma_2 < 1$, 计算 $f(x_0)$, 设置 $k=0$;

②函数近似:定义范数 $\|\cdot\|_k$, 定义信任域 $B_k = \{x \in R^n \mid \|x - x_k\|_k \leq \Delta_k\}$, 在此信任域内寻找 $f(x)$ 的近似函数 m_k ;

^{*} 国防基础预研项目基金、航天创新基金和航空科学基金项目(No.02153073)。贾静平 博士研究生。

③信任域子问题:计算增量 s_k 使得 m_k 取得最小并且 $x_k + s_k \in B_k$;

④信任域保真度:

计算 $f(x_k + s_k)$, 和 $r_k = \frac{f(x_k) - f(x_k + s_k)}{m_k(x_k) - m_k(x_k + s_k)}$, 如果 $r_k \geq \eta$, 那么增量 s_k 可接受, $x_{k+1} = x_k + s_k$; 否则 $x_{k+1} = x_k$, 因为 η 是一个小的正数, 这个条件确保 s_k 使得目标函数 f 取值减小;

⑤更新信任域半径:

$$\Delta_{k+1} \in \begin{cases} [\Delta_k, \infty) & \text{如果 } r_k \geq \eta \\ [\gamma_2 \Delta_k, \Delta_k) & \text{如果 } r_k \in [\eta, \eta_2) \\ [\gamma_1 \Delta_k, \gamma_2 \Delta_k) & \text{如果 } r_k < \eta \end{cases}$$

, 即当 m_k 很好地近似了 f 时, r_k 会取得较大的值, 当 $r_k \geq \eta$ 时, 就扩展信任域半径。如果 $r_k < \eta$ 或甚至为负, 就说明在 B_k 内, m_k 不是很好近似了 f , 那么就不接受 s_k , 并且缩小信任域半径, 在更小的邻域内寻找 f 的近似;

⑥如果 Δ_k 小于预先设定的阈值 Δ_{end} 或 k 超过最大循环次数 MAX_{iter} , 退出; 否则 $k = k + 1$, 跳至步骤 2。

1.2 改进的 QP-TR 信任域算法

与经典信任域算法相比, QP-TR 信任域算法^[2]改进了近似函数 m_k 的获得方式。在 B_k 中用 n 维空间中的二次多项式 m_k 来近似目标函数, 即:

$$m_k(x_k + s) = m_k(x_k) + \langle g_k, s \rangle + \frac{1}{2} \langle s, H_k s \rangle, m_k(x_k) = f(x_k)$$

m_k 通过多变量拉格朗日插值法得到。目标函数在 x_k 处的海森矩阵 H_k 不是通过 BFGS 方法多次循环近似估计的, 而是

直接从 m_k 计算出来, 从而避免了 BFGS 方法在目标函数 f 曲率快速变化时精度差的缺点。 $g_k = \nabla_x m_k(x_k)$ 也是直接从 m_k 得到, 避免了使用直接差分计算导数而引入的误差。因此, 信任域子问题求解的关键就在于寻找增量 s_k , 并使得 $s_k = \arg \min_{\|s\|_k \leq \Delta_k} \Psi_k(s) \equiv \langle g_k, s \rangle + \frac{1}{2} \langle s, H_k s \rangle$ 。QP-TR 的其余步骤与经典信任域方法类似。

2 结合尺度空间理论的跟踪算法

2.1 目标特征

HSV 空间能有效地将色调(H)、饱和度(S)和亮度(V)区分开来, 使得用色调作为目标特征, 对光照变化具有很好的鲁棒性^[7]。首先, 计算目标区域 HSV 颜色空间中 H 通道的直方图(注意, 当亮度值太小时, 饱和度值也很小, 色调值就特别容易受到噪声干扰; 而当亮度值太大时, 多数像素的色调偏向白色, 不利于特征的提取。因此在计算直方图时, 我们通常忽略亮度值太小或太大的像素)。然后, 在跟踪过程中, 对每一帧输入图像使用上述直方图作为查找表, 计算图像中每个像素归属目标的概率, 从而得到目标的概率分布图, 取值范围为 $[0, 0, 1, 0]$ 。最后, 为便于后续处理和显示, 本文将将其规整到 $[0, 255]$ 。图 1 给出了生成概率分布图的例子。左图中红色方框代表目标区域。首先计算该目标的颜色直方图, 亮度取值范围为 $[50, 250]$, 然后利用该直方图计算右图的目标概率分布图, 其结果如中图所示。可以看出, 与目标颜色接近的像素在概率分布图中有较大的概率值。



图 1 目标概率分布图的计算

为了精确跟踪存在缩放的目标, 我们使用 Lindeberg 的尺度空间理论^[5]作为最佳尺度选择的依据。给定一个 D 维空间中的连续函数 $f: R^D \rightarrow R$ 和 $D+1$ 维的高斯核 $g: R^D \times R_+ \rightarrow R$, $g(x; t) = \frac{1}{(2\pi t)^{D/2}} e^{-(x_1^2 + \dots + x_D^2)/(2t)}$, 那么该函数的线性尺度空间表示 $L: R^D \times R_+ \rightarrow R$ 可被定义为 f 和 g 的卷积 $L(\cdot; t) * f(\cdot)$, 具有可变宽度 t 。 L 的规范化导数具有尺度不变性, 说明如下:

如果两个函数 f 和 $\tilde{f}$ 有 $f(x) = \tilde{f}(sx)$ 的关系, L 和 $\tilde{L}$ 分别是它们的尺度空间表示,

$$\begin{aligned} L(\cdot; t) &= g(\cdot; t) * f(\cdot) \\ \tilde{L}(\cdot; \tilde{t}) &= g(\cdot; \tilde{t}) * \tilde{f}(\cdot) \end{aligned} \quad (1)$$

那么空间域变量和尺度参数有如下关系:

$$\begin{aligned} \tilde{x} &= sx \\ \tilde{t} &= s^2 t \end{aligned} \quad (2)$$

并且有 $L(x; t) = \tilde{L}(\tilde{x}; \tilde{t})$ 。第 m 阶空间域微分满足 $\partial_x^m L(x; t) = s^m \partial_{\tilde{x}}^m \tilde{L}(\tilde{x}; \tilde{t})$ 。作变量代换 $\xi = \frac{x}{t^{1/2}}$, 则有:

$$\partial_{\xi} = t^{1/2} \partial_x$$

$$\partial_{\tilde{\xi}} = \tilde{t}^{1/2} \partial_{\tilde{x}} \quad (3)$$

进而有

$$\partial_{\tilde{x}}^m L(x; t) = s^m (1-\gamma) \partial_{\tilde{x}}^m \tilde{L}(\tilde{x}; \tilde{t}) \quad (4)$$

$\partial_{\tilde{x}}^m L(x; t)$ 称为 γ 规范化导数。当 $\gamma=1$ 时, $\partial_{\tilde{x}}^m L(x; t) = \partial_{\tilde{x}}^m \tilde{L}(\tilde{x}; \tilde{t})$, 即此时规范化导数的取值具有尺度不变性。这个良好性质说明, 如果在 f 的尺度空间中的 $(x_0; t_0)$ 点, γ 规范化导数取得了最大值, 那么在 $\tilde{f}$ 的尺度空间中的 $(s x_0; s^2 t_0)$, 也会有一个对应的最大值。利用这一优良的性质, 只要适当选取 γ 规范化导数的表达式, 我们就可以检测函数的尺度。

灰度图像可以看成是二维函数 f , 即 $D=2$, $f: R^2 \rightarrow R$ 。当 $\gamma=1$ 时, $(t^2 \nabla^2 L)^2 = (t \nabla^2 L)^2 = (t(L_{xx} + L_{yy}))^2$ 反映了图像中灰度块(blob)的分布情况。我们定义 $\varphi(x, y, t) = (t(L_{xx}(x, y) + L_{yy}(x, y)))^2$ 为图像的多尺度规范化 Laplacian 滤波函数。在不同的尺度下, $\varphi(x, y, t)$ 在不同大小的灰度块处取得极大值。图 2 说明了这一情况, 在原灰度图像中包含大小不同的两个灰度块, 当 $t=90$ 时, $\varphi(x, y, t)$ 在小块的中心处取得极大值, 当 $t=391$ 时, $\varphi(x, y, t)$ 在大块的中心处取得极大值。所以, 通过检测 $\varphi(x, y, t)$ 在不同尺度下不同位置处的极值就可以定位图像中的灰度块, 并能决定它们的尺度。

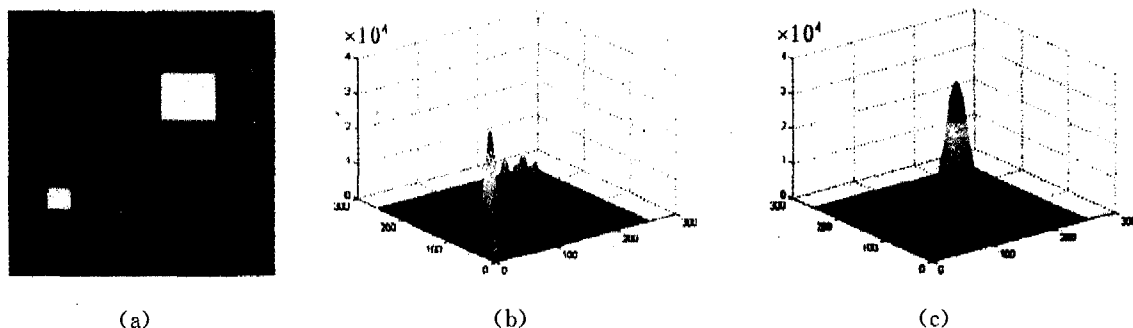


图2 图像中大小不同的灰度块在不同尺度下 $\varphi(x, y, t)$ 在不同位置取得极大值
(a)原图; (b) $t=90$ 时, $\varphi(x, y, t)$ 的取值; (c) $t=391$ 时, $\varphi(x, y, t)$ 的取值。

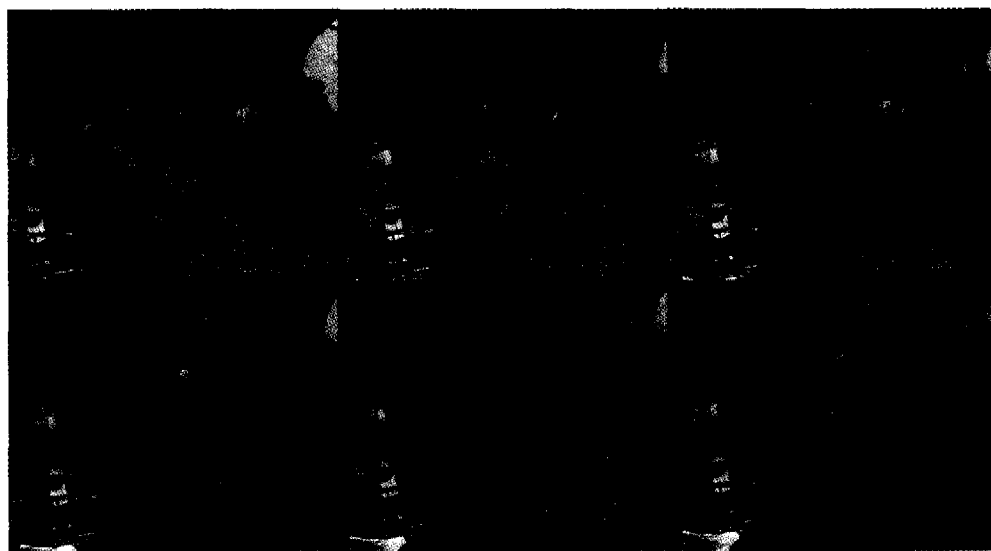


图3 “骑车人”序列跟踪结果
从左至右, 从上至下依次为第 20, 56, 63, 111, 121, 132 帧。

在使用目标直方图计算得到的目标概率分布图中, 每个像素值代表该像素属于目标的概率, 目标在该图中对应着一个灰度块。真实的目标位置和正确的目标尺度, 对应着该灰度块在概率分布图 $\varphi(x, y, t)$ 滤波中的一个极大值, 通过 QP-TR 算法可以在连续的尺度空间中寻找这个极大值, 并确定相应的目标真实位置和尺度。因此将颜色直方图作为目标特征时, 可将 $-\varphi(x, y, t)$ 作为目标函数, 用 QP-TR 算法求解。

2.2 跟踪算法

使用颜色直方图作为目标特征, 结合 QP-TR 方法进行优化的跟踪算法适用于目标存在缩放的情况, 其步骤如下:

1) 初始化, 选定目标, 计算目标方框内的 HSV 空间 H 通道直方图作为目标特征;

2) 对输入的每一帧图像, 进行缩放, 使得缩放后的图像中目标的宽高比为 1:1。若目标的宽、高分别为 w_T 和 h_T , 帧的宽、高分别为 w 和 h , 缩放后的图像的宽高比为 $\rho = wh_T/h/w_T$ 。若保持图像宽度不变, 则缩放后图像宽度应变为原宽度的 $w/h/\rho$ 倍;

3) 对缩放后的图像, 按照目标直方图计算概率分布图。初始化三维向量 $x_0 = (x_{prev}, y_{prev}, t_{prev})^T$, 其中 (x_{prev}, y_{prev}) 是上一帧中目标中心点的位置, t_{prev} 是上一帧中目标的尺度参数。初始化信任域半径初始值 $\Delta_0 = 9$, 中止值 $\Delta_{end} = 0.1$, 最大循环次数 $MAX_{iter} = 1000$;

4) 以 $f(x) = -\varphi(x, y, t)$ 作为目标函数使用 QP-TR 信

任域算法优化, 优化得到 $x_{opt} = (x_{opt}, y_{opt}, t_{opt})^T$, (x_{opt}, y_{opt}) 即为目标在本帧中中心点位置, t_{opt} 为本帧中目标的尺度参数, 转 2)。

3 结合灰度模板的跟踪算法

3.1 目标特征

将选定目标区域内所有像素灰度值的集合作为目标特征, 这个集合就称为特征模板。若目标模板为 W , 大小为 $M \times N$, 针对一帧中任一点 (u, v) , 定义平均绝对误差^[8]为 $MAD(u, v) = \frac{1}{M \cdot N} \sum_{0 \leq m < M, 0 \leq n < N} |W(m, n) - I(u+m, v+n)|$ 。那么最小的 MAD 就对应着目标的正确位置。因此目标跟踪问题转化成了优化问题, 即寻找一点 (u, v) 使得 $(u, v) = \arg \min_{(u, v)} MAD(u, v)$, 目标函数为 $MAD(u, v)$, 可以使用 QP-TR 方法求解。

3.2 跟踪算法

使用灰度模板作为目标特征, 结合 QP-TR 方法进行优化的跟踪算法适用于实时目标跟踪, 其步骤如下:

1) 初始化, 选定目标, 保存目标所在矩形区域的灰度值作为目标模板;

2) 初始化二维向量 $x_0 = (u_{prev}, v_{prev})^T$, 其中 (u_{prev}, v_{prev}) 是上一帧中目标左上角点的位置, 初始化信任域半径初始值 $\Delta_0 = 20$, 中止值 $\Delta_{end} = 0.5$, 最大循环次数 $MAX_{iter} = 1000$;

3) 以 $f(x) = MAD(u, v)$ 作为目标函数使用 QP-TR 信任域算法优化, 优化得到 $x_{opt} = (u_{opt}, v_{opt})^T$, (u_{opt}, v_{opt}) 即为目标在本帧中左上角点位置, 转 2)。

4 跟踪实例

为了验证本文算法的性能, 我们在真实序列上进行了大量的实验, 并将本文提出的两种算法和已有的算法进行比较。图 3 是使用第 3 节提出的基于灰度模板的算法对“骑车人”序列进行跟踪的结果。

我们算法的搜索窗为 20×20 ($\Delta_0 = 20$), 比三步搜索法^[8]的 15×15 大, 但在相同精度下每帧搜索块的次数却明显比三步搜索法少。图 4 说明了这一情况, 图中实线 (QP-TR) 为本文算法在各帧中的搜索次数, 虚线 (TSS) 为三步搜索方法的搜索次数。本文算法平均为每帧 19.85 次, 少于三步搜索法的每帧固定 25 次。

图 5 是在“行人”序列上, 使用第 2 节尺度空间算法与带宽矩阵 Mean Shift 算法^[4]对比的跟踪结果。图中第二行是带宽矩阵 Mean Shift 算法的结果, 该算法在水平和垂直两个方向上试探比原尺寸大和小 10% 的两个尺寸以猜测目标的实际尺寸, 其尺度参数只能取三个离散值, 因此在跟踪过程中对

目标大小刻画有较大的误差, 导致跟踪精度下降 (如 217、317 帧)。第一行是本文的算法的结果, 由于是在连续的尺度空间里决定目标的当前尺度, 因此目标方框随目标缩小而连续缩小, 精确描述了目标的大小变化。

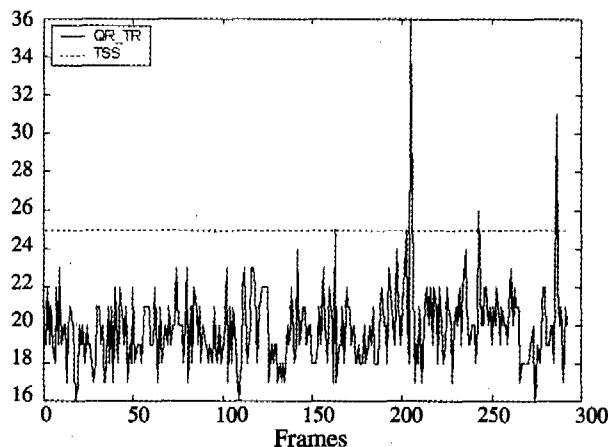


图 4 本文算法和三步搜索法每帧搜索次数比较

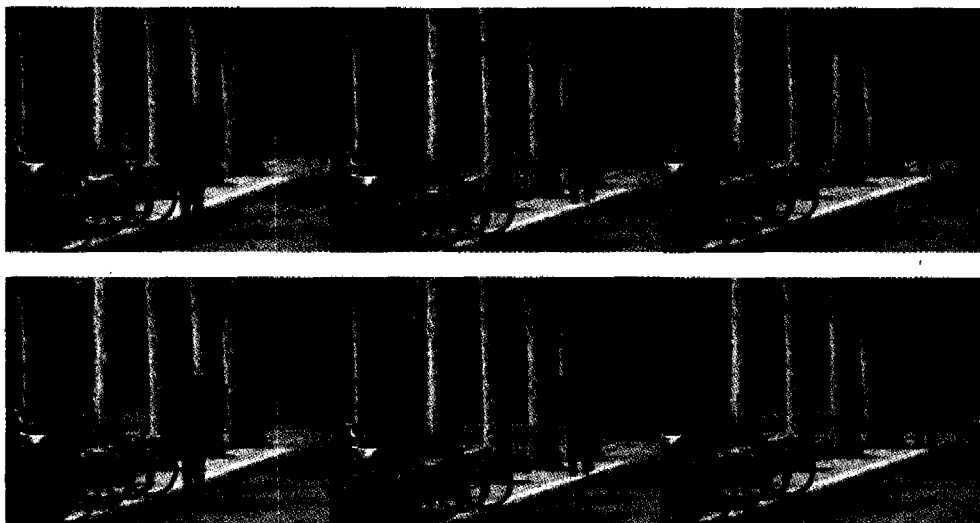


图 5 “行人”序列上和带宽矩阵 Mean Shift 跟踪算法^[4]比较
上列: 尺度空间算法, 下列: 文[4]算法, 从左至右为第 3、217 和 317 帧。

结束语 本文将信任域算法和尺度空间理论相结合, 成功解决了视频跟踪中正确选取目标尺度的问题, 将信任域算法与灰度模板相结合, 解决了扩大搜索窗同时减少计算量的问题。从 Lindeberg 尺度空间理论开始, 我们首先引入了使用多尺度规范化 Laplacian 滤波函数来检测灰度块的方法, 然后通过 QP-TR 信任域算法在图像平面和尺度空间中搜索多尺度规范化 Laplacian 滤波函数的极大值, 使每个极大值都对应着一个灰度块在图像平面内的位置和尺度, 从而实现了目标的精确定位和跟踪。本文利用目标的颜色特征将原始图像转换成目标概率分布图, 从原理上来说尺度空间是目标概率分布图通过与一族空间 LOG 滤波器卷积得到的。显式地产生这个尺度空间不仅计算量大, 而且只能取有限个离散的尺度值, 而 QP-TR 信任域算法有效地解决了在连续的尺度空间快速搜索极值的问题。在真实序列上进行的实验结果表明, 我们的算法不仅能自动选取最佳的目标尺度, 而且与已有算法相比, 明显提高了跟踪的精度和速度。

参考文献

- Comaniciu D, Ramesh V, Meer P. Kernel-Based Object Tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 564~577
- Berghen F V. Intermediate Report on the development of an optimization code for smooth, continuous objective functions when derivatives are not available. http://www.optimization-online.org/DB_HTML/2003/08/704.html
- Comaniciu D, Ramesh V, Meer P. Real-Time Tracking of Non-Rigid Objects Using Mean Shift. *IEEE Computer Vision and Pattern Recognition*, 2000, 2: 142~149
- Jia Jingping, Zhao Rongchun. Tracking of objects in image sequences using bandwidth matrix mean shift algorithm. In: the 7th International Conference on Signal Processing Proceedings, 2004, 2: 918~921
- Lindeberg T. Feature Detection with Automatic Scale Selection. *International Journal of Computer Vision*, 1998, 30(2): 79~116
- Liu Tyng-Luh, Chen Hwang-Tzong. Real-Time Tracking Using Trust-Region Methods. *IEEE PAMI*, 2004, 26(3): 397~402
- Bradski G R. Computer Vision Face Tracking for Use in a Perceptual User Interface. In: *IEEE Workshop on Applications of Computer Vision*, Princeton, NJ, 1998, 214~219
- 任金昌, 张文哲, 等. 一种基于自适应阈值的复杂背景下自动目标跟踪方法. *计算机应用研究*, 2003, 20(4): 55~57