



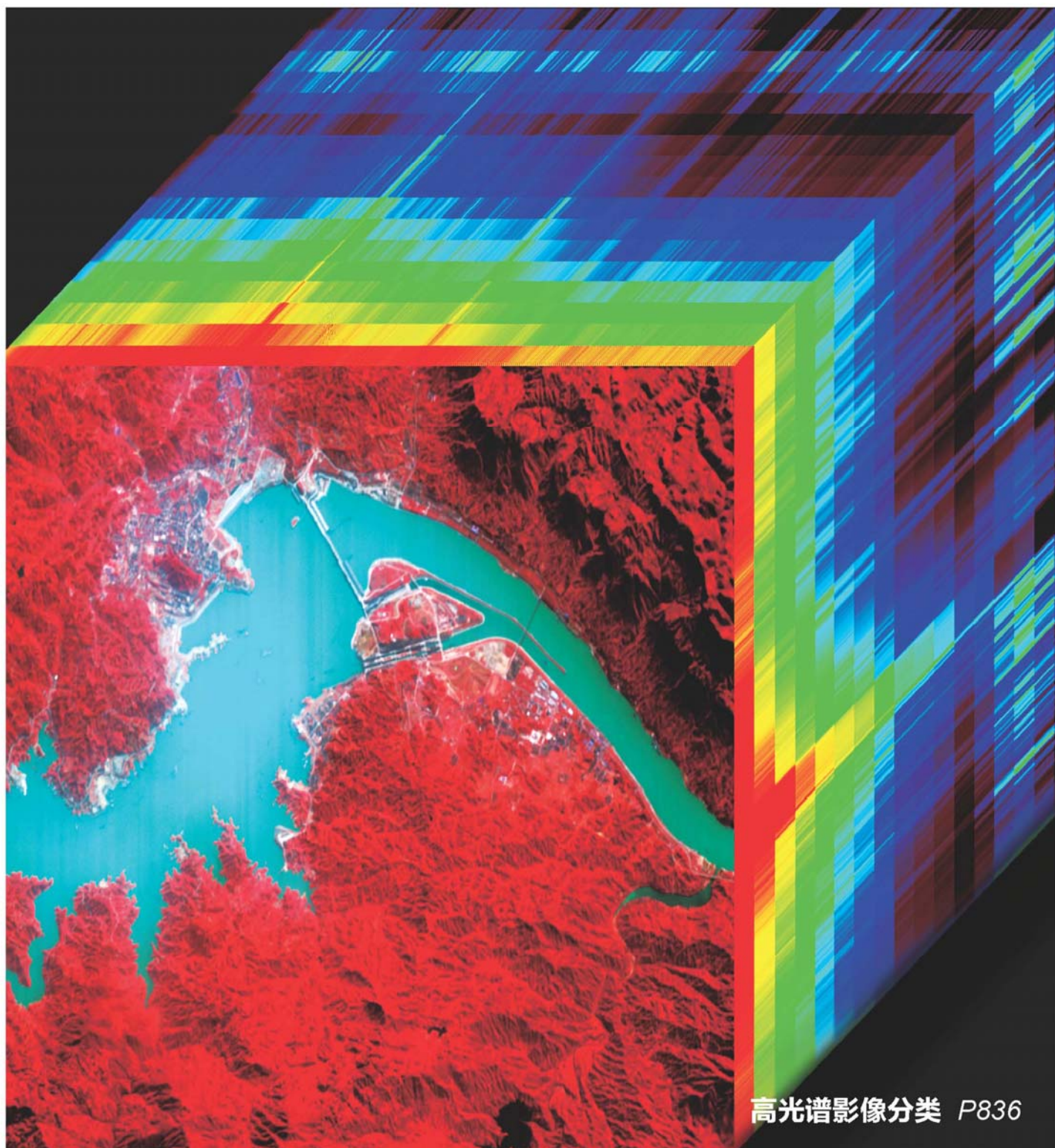
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

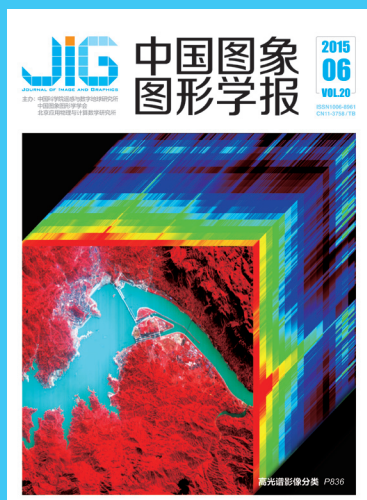
中国图象图形学报

2015
06
VOL.20

ISSN1006-8961
CN11-3758/TB



高光谱影像分类 P836



第20卷第6期（总第230期）

2015年6月16日

中国精品科技期刊
中国期刊影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 李小文

编辑出版《中国图象图形学报》编辑出版委员会

邮政信箱 北京9718信箱

邮 编 100101

电子信箱 jig@radi.ac.cn

电 话 010-64807995

网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

（邮政信箱：北京399信箱 邮编：100048）

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen

Editor, Publisher Editorial and Publishing Board of Journal of

Image and Graphics

P.O.Box 9718, Beijing, P.R.China

Zip code 100101

E-mail jig@radi.ac.cn

Telephone 010-64807995

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation

(P.O.Box 399, Beijing 100048, P.R.China))

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

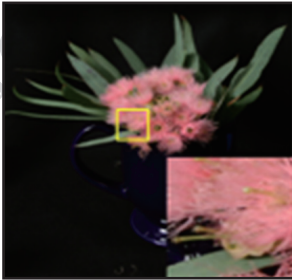
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

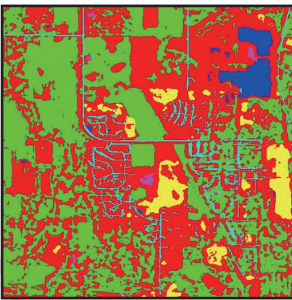
国内定价 50.00元



边缘结构保持的加权BDTV
全光场图像超分辨率重建
(第0733页)



基于超像素的多主体图像交
互分割(第0764页)



谐波分析光谱角制图高光谱
影像分类(第0836页)

图像处理和编码

边缘结构保持的加权BDTV全光场图像超分辨率重建

张旭东, 李梦娜, 张骏, 胡良梅, 王一 0733

基于Arnold映射的分块双层自适应扩散图像加密算法

徐亚, 张绍武 0740

结合梯度信息的特征相似性图像质量评估

苗莹, 易三莉, 贺建峰, 邵党国 0749

图像分析和识别

复杂场景下自适应背景减除算法

邵奇可, 周宇, 李路, 陈庆章 0756

基于超像素的多主体图像交互分割

伯彭波, 袁野, 王宽全 0764

采用金字塔纹理和边缘特征的图像烟雾检测

李红娣, 袁非牛 0772

结合手指检测和HOG特征的分层静态手势识别

刘淑萍, 刘羽, 於俊, 汪增福 0781

形状角计算改进及其在矩形目标识别中的应用

孟伟荣, 李辉, 方黎勇, 白金平 0789

融合图像感知哈希技术的运动目标跟踪

李子印, 朱明凌, 陈柱 0795

利用快速傅里叶变换的目标尺度自适应回归跟踪

张浪, 侯志强, 余旺盛, 许婉君 0805

基于点云属性信息的平面靶特征提取

刘燕萍, 贾东峰 0815

融合局部思想和协作表达的鲁棒分类

楼宋江, 赵小明, 于海涛, 张石清 0822

计算机图形学

草绘3维人体建模的模板形变方法

李毅, 刘兴川, 孙亭 0828

遥感图像处理

谐波分析光谱角制图高光谱影像分类

杨可明, 刘飞, 孙阳阳, 魏华锋, 史钢强 0836

倾斜航空影像的城区DSM生成

吴军, 程门门, 姚泽鑫, 彭智勇, 李俊, 马峻 0845

封面图片由中国科学院遥感与数字地球研究所高光谱遥感应用技术研究室张霞研究员提供

CONTENTS

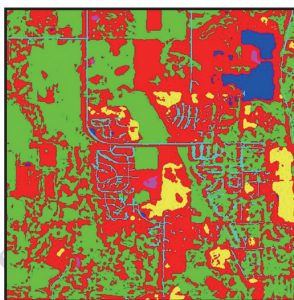
JOURNAL OF IMAGE AND GRAPHICS



Edge preserved light field image super-resolution based on weighted BDTV model (P0733)



Interactive multiphase image segmentation based on superpixels (P0764)



Classification algorithm of hyperspectral imagery by harmonic analysis and spectral angle mapping (P0836)

Image Processing and Coding

Edge preserved light field image super-resolution method based on weighted BDTV model

Zhang Xudong, Li Mengna, Zhang Jun, Hu Liangmei, Wang Yi 0733

Encryption algorithm of image blocking and double adaptive diffusion with Arnold mapping

Xu Ya, Zhang Shaowu 0740

Image quality assessment of feature similarity combined with gradient information

Miao Ying, Yi Sanli, He Jianfeng, Shao Dangguo 0749

Image Analysis and Recognition

Adaptive background subtraction approach of Gaussian mixture model

Shao Qike, Zhou Yu, Li Lu, Chen Qingzhang 0756

Interactive multiphase image segmentation based on superpixels

Bo Pengbo, Yuan Ye, Wang Kuanquan 0764

Image based smoke detection using pyramid texture and edge features

Li Hongdi, Yuan Feiniu 0772

Hierarchical static hand gesture recognition by combining finger detection and HOG features

Liu Shuping, Liu Yu, Yu Jun, Wang Zengfu 0781

Improved shape angle computation and its application in detection of rectangular

Meng Weirong, Li Hui, Fang Liyong, Bai Jinping 0789

Object tracking algorithm based on perception hash technology

Li Ziyin, Zhu Mingling, Chen Zhu 0795

Scale adaptive regression tracking method based on fast Fourier transform

Zhang Lang, Hou Zhiqiang, Yu Wangsheng, Xu Wanjuan 0805

Feature extraction of flat target based on point cloud attribute information

Liu Yanping, Jia Dongfeng 0815

Locality constrained collaborative representation classification for robust classification

Lou Songjiang, Zhao Xiaoming, Yu Haitao, Zhang Shiqing 0822

Computer Graphics

Sketch-based method for 3D human modeling using template deformation

Li Yi, Liu Xingchuan, Sun Ting 0828

Remote Sensing Image Processing

Classification algorithm of hyperspectral imagery by harmonic analysis and spectral angle mapping

Yang Keming, Liu Fei, Sun Yangyang, Wei Huafeng, Shi Gangqiang 0836

Automatic generation of high-quality urban DSM with airborne oblique images

Wu Jun, Cheng Menmen, Yao Zexin, Peng Zhiyong, Li Jun, Ma Jun 0845

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2015)06-0805-10

论文引用格式: Zhang L, Hou Z Q, Yu W S, Xu W J. Scale adaptive regression tracking method based on fast Fourier transform[J]. Journal of Image and Graphics, 2015, 20(6): 0805-0814. [张浪, 侯志强, 余旺盛, 许婉君. 利用快速傅里叶变换的目标尺度自适应回归跟踪[J]. 中国图象图形学报, 2015, 20(6): 0805-0814.] [DOI:10.11834/jig.20150610]

利用快速傅里叶变换的目标尺度自适应回归跟踪

张浪, 侯志强, 余旺盛, 许婉君

空军工程大学信息与导航学院, 西安 710077

摘要: 目的 视频跟踪中, 跟踪背景复杂及目标表观变化是导致跟踪失败的主要原因。回归跟踪算法利用目标的表观信息建立回归模型进行跟踪, 然而该算法的跟踪效率较低; 基于循环结构的检测跟踪方法能有效利用循环结构提高跟踪效率, 但该算法不能跟踪尺度变化的目标。为解决上述问题, 本文提出一种基于快速傅里叶变换的尺度自适应回归跟踪算法, 方法 首先利用快速傅里叶变换建立目标的核岭回归模型并搜索得到目标的中心位置, 然后计算候选目标区域像素点的权重分布图, 从而估计出目标的最佳尺度。结果 进行6组实验, 与当前常见算法(CBWH、IVT、DFT、DSST、增量试探法)相比, 本文算法不仅能较好地适应背景、目标尺度及姿态的变化, 而且平均每帧运行时间短(毫秒级)。结论 本文提出一种基于快速傅里叶变换的尺度自适应回归跟踪算法, 算法对背景、尺度及姿态变化的目标跟踪具有较强的鲁棒性和很好的跟踪效率。

关键词: 视频跟踪; 尺度自适应; 核岭回归; 快速傅里叶变换

Scale adaptive regression tracking method based on fast Fourier transform

Zhang Lang, Hou Zhiqiang, Yu Wangsheng, Xu Wanjun

The Information and Navigation Institute of Air Force Engineering University, Xi'an 710077, China

Abstract: **Objective** Complicated tracking environments and changes in object appearance are the major causes of failure in visual tracking. The regression tracking algorithm facilitates tracking by establishing a regression model on the basis of information regarding an object's appearance. However, this algorithm displays low tracking efficiency. The use of a tracking-by-detection algorithm that is based on circular structures can improve this efficiency, but such an algorithm is unsuited for handling changes in the scale of an object. Therefore, a scale-adaptive regression tracking algorithm based on fast Fourier transform is proposed to address these problems. **Method** First, the algorithm determines the center position of the object in the search region using the regression model. The algorithm then estimates the ideal scale by considering the weight image of all pixels in the candidate region. **Result** Comparing with the popular algorithm such as CBWH, IVT and so on, the results from six experiments indicate that the proposed algorithm can not only adapt to changes in background, object scale, and pose, but that the running time of each frame is short as well. **Conclusion** This paper presents a scale-adaptive regression tracking algorithm based on fast Fourier transform. It owns good robustness and efficiency for tracking target with the changing of background, object's scale and pose.

Key words: visual tracking; scale adaptation; kernel ridge regression; fast Fourier transform

收稿日期: 2014-11-27; 修回日期: 2015-02-12

基金项目: 国家自然科学基金项目(61175029, 61473309)

第一作者简介: 张浪(1990—), 男, 现为空军工程大学信号与信息处理专业在读硕士研究生, 主要研究方向为视觉目标跟踪。

E-mail: zhanglangwy@126.com

Supported by: National Natural Science Foundation of China (61175029, 61473309)

0 引言

目标跟踪问题是当前计算机视觉领域中的一个热点问题,由于广泛应用于视频监控、图像压缩、3维重构等领域,其受到越来越多学者的关注^[1-2]。但是,复杂背景的干扰以及目标表观的变化往往造成跟踪失败,因此,适应跟踪背景及目标表观变化是目标跟踪中的难点问题。

近些年来,出现了大量的跟踪算法。Mean Shift算法^[3-4]是一种典型的核跟踪算法,该方法采用目标直方图特征进行跟踪,由于忽略了目标的空间结构信息,容易匹配到具有相似直方图特征的背景中去。文献[5]通过目标的正负样本集建立目标的稀疏^[6]相似图,然后利用该相似图搜索与目标最相似的候选目标,但当背景干扰严重时,算法不能正确搜索到最佳候选目标。文献[7]采用分布场对目标进行描述,该方法能减轻模型对于空间结构信息的依赖度,然而当目标尺度及姿态变化较大时,算法容易跟踪失败。文献[8]采用修正的背景加权直方图方法对目标进行建模,该方法能降低背景的干扰,但在背景发生明显变化时跟踪效果不佳。上述算法由于复杂背景的干扰或者目标表观的变化,造成了跟踪失败。

基于循环结构的检测跟踪法^[9]提出利用循环结构矩阵建立目标回归模型进行跟踪,该方法不仅利用了目标表观特征,能一定程度上适应复杂背景的干扰,而且通过构造循环结构矩阵能有效提高跟踪效率。然而当目标尺度发生较大变化时,目标模型不能充分描述目标的表观特征,因而容易跟踪失败。受此启发,本文提出一种目标尺度自适应的回归跟踪方法,该方法跟踪过程主要有两个步骤:首先建立目标的回归模型,并利用快速傅里叶变换快速求解模型,然后在文献[10]算法原理基础上提出了一种利用候选目标区域相似性权重分布图快速估计出目标最佳尺度的方法。通过跟踪实验表明:本文算法不仅能较好地适应背景、目标尺度及姿态的变化,具有较强的鲁棒性,而且具有很高的跟踪效率。

1 基于快速傅里叶变换的回归模型

核岭回归(KRR)是一种以结构风险最小化为

学习准则,解决在原始样本空间中不能用线性方法求解的非线性问题的算法,该算法具有高的泛化能力。本节首先分析线性岭回归算法及其局限性,然后引入非线性核岭回归算法^[11](KRR)。

1.1 核岭回归

线性回归是利用数理统计中的回归分析,可以用来确定两个或两个以上变量之间映射关系的一种统计分析方法,输出变量是输入变量线性组合,即

$$\mathbf{y} = \mathbf{w}^T \mathbf{x} + \zeta \quad (1)$$

式中, $\mathbf{w}^T \mathbf{x}$ 是向量之间的内积, ζ 表示预测值与真实值的偏差,一般服从标准正态分布 $\zeta \propto N(0, \delta^2)$ 。

对于不适定问题,最小二乘法的解是病态的,因此,Hoerl 和 Kennard 提出代价函数

$$J(\mathbf{w}) = \sum_{i=1}^N (y_i - \mathbf{w}^T \mathbf{x}_i - \zeta)^2 + \lambda \|\mathbf{w}\|_2^2 \quad (2)$$

式中, $\|\mathbf{w}\|_2^2 = \mathbf{w}^T \mathbf{w}$ 为正则化项, λ 为正则化参数, λ 在最小正则化项和最小风险平方和之间起平衡作用。得到岭回归解

$$\hat{\mathbf{w}}_{\text{ridge}} = (\lambda \mathbf{I}_N + \mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T \mathbf{y} \quad (3)$$

式中, $\mathbf{x} \in \mathbf{R}^{D \times N}$, $\mathbf{y} \in \mathbf{R}^N$, $\mathbf{I}_N$ 是 N 行 N 列的单位矩阵。根据矩阵求逆引理,可得原始解的对偶解

$$\hat{\mathbf{w}}_{\text{dual}} = \mathbf{x}^T (\lambda \mathbf{I}_N + \mathbf{x} \mathbf{x}^T)^{-1} \mathbf{y} \quad (4)$$

但是岭回归对于非线性问题具有很大的局限性。利用核思想,可将线性回归转化为非线性回归。将内积 $\mathbf{x} \mathbf{x}^T$ 用核矩阵 $\mathbf{K}$ 代替,得到核岭回归的解

$$\hat{\mathbf{w}}_{\text{kernel}} = \mathbf{x}^T (\lambda \mathbf{I}_N + \mathbf{K})^{-1} \mathbf{y} \quad (5)$$

必须注意,此处 $\mathbf{x}$ 为向高维映射后的变量,不能直接进行运算,为方便计算,后面将会引入变量。

1.2 核空间的循环卷积

假设有两个向量 $\mathbf{p} \in \mathbf{R}^M$ 和 $\mathbf{q} \in \mathbf{R}^N$, 其卷积结果为一个 $M+N-1$ 的向量,即

$$\mathbf{p} * \mathbf{q} = \sum_i p_i q_{\tau-i} \quad (6)$$

通过构造循环矩阵,将卷积转化为乘积的形式,即

$$\mathbf{p}' = \{\underbrace{p_1, p_1, \dots, p_1}_M, \underbrace{0, 0, \dots, 0}_{N-1}\} \quad (7)$$

$$\mathbf{q}' = \{\underbrace{q_1, q_1, \dots, q_1}_N, \underbrace{0, 0, \dots, 0}_{M-1}\} \quad (8)$$

$$C(\mathbf{p}') \mathbf{q}'^T = \begin{bmatrix} p'_1 & p'_{M+N-1} & \cdots & p'_2 \\ p'_2 & p'_1 & \cdots & p'_3 \\ \vdots & \vdots & & \vdots \\ p'_{M+N-1} & p'_{M+N-2} & \cdots & p'_1 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ 0 \end{bmatrix} \quad (9)$$

根据卷积定理,可得

$$C(\mathbf{p}')\mathbf{q}' = F^{-1}(F(\mathbf{p}')F(\mathbf{q}')) \quad (10)$$

由于一幅图像对应一个像素矩阵,则对于一个图像块 X ,可以将其转化为一个 n 维向量 $\mathbf{x}$ 并将其扩展为循环矩阵

$$C(\mathbf{x}) = \begin{bmatrix} \mathbf{E}_0\mathbf{x} \\ \mathbf{E}_1\mathbf{x} \\ \vdots \\ \mathbf{E}_{n-1}\mathbf{x} \end{bmatrix}^T \quad (11)$$

式中, $\mathbf{E}_i$ 为置换矩阵,用 $\mathbf{E}_i$ 与 $\mathbf{x}$ 相乘可以将 $\mathbf{x}$ 中元素进行任意交换,即

$$\mathbf{E}_i = \begin{bmatrix} \overbrace{0 \cdots 0}^i & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & \cdots & 1 & 0 \end{bmatrix}_{n \times n} \quad (12)$$

由循环矩阵 $C(\mathbf{x})$ 中每行构成的样本,形成的核矩阵 $\mathbf{K}_{ij} = \kappa(\mathbf{E}_i\mathbf{x}, \mathbf{E}_j\mathbf{x})$ 是循环矩阵^[9]。采用高斯核函数,则核空间下的卷积可表示为

$$\mathbf{k}_i^{\text{gauss}} = \kappa(\mathbf{x}', \mathbf{E}_i\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x}'\|^2 + \|\mathbf{x}\|^2 - 2\mathbf{x}'^T\mathbf{E}_i\mathbf{x}}{2\delta^2}\right) \quad (13)$$

由于 $\mathbf{E}_i$ 是置换矩阵,因此核函数可写成

$$\begin{aligned} \mathbf{k}^{\text{gauss}} &= \exp\left(-\frac{\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 - 2\mathbf{C}(\mathbf{x}')\mathbf{x}}{2\delta^2}\right) = \\ &\exp\left(-\frac{\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 - 2F^{-1}(F(\mathbf{x}')F(\mathbf{x}))}{2\delta^2}\right) \end{aligned} \quad (14)$$

2 目标最佳尺度估计方法

2.1 Mean Shift 目标跟踪原理

Mean Shift 算法是一种经典的核跟踪方法,适用于目标可以表示成矩形和椭圆形的目标跟踪。该算法用目标的颜色直方图对目标进行描述。

目标直方图模型为

$$\mathbf{q}_u = C \sum_{i=1}^n k(\|\mathbf{x}_i^*\|^2) \delta[b(\mathbf{x}_i^*) - u] \quad (15)$$

候选目标直方图模型为

$$\mathbf{p}_u(y) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - \mathbf{x}_i}{h}\right\|^2\right) \delta[b(\mathbf{x}_i) - u] \quad (16)$$

式中, C 和 C_h 为归一化常数, h 为带宽参数, $\mathbf{x}_i^*$ 和 $\mathbf{x}_i$ 分别为目标区域和候选目标区域的特征位置。 $k(\mathbf{x})$ 为非负、递减、分段连续的核函数。

为描述目标直方图模型和候选目标直方图模型的相似度,采用 Bhattacharyya 系数 $\rho(y)$ 度量,即

$$\rho(y) \equiv \rho[\mathbf{p}_u(y), \mathbf{q}_u] = \sum_{u=1}^m \sqrt{\mathbf{p}_u(y) \mathbf{q}_u} \quad (17)$$

运用 Taylor 扩展, Bhattacharyya 系数线性近似为

$$\begin{aligned} \rho[\mathbf{p}_u(y), \mathbf{q}_u] &\approx \\ &\frac{1}{2} \sum_{u=1}^m \sqrt{\mathbf{p}_u(y) \mathbf{q}_u} + \frac{C_h}{2} \sum_{u=1}^{n_h} w_i k\left(\left\|\frac{y - \mathbf{x}_i}{h}\right\|^2\right) \end{aligned} \quad (18)$$

在 Mean Shift 迭代中,目标从当前位置 y_0 移动到新位置 y_1 ,即

$$y_1 = \frac{\sum_{i=1}^{n_h} \mathbf{x}_i w_i g(\|\mathbf{y}_0 - \mathbf{x}_i/h\|^2)}{\sum_{i=1}^{n_h} w_i g(\|\mathbf{y}_0 - \mathbf{x}_i/h\|^2)} \quad (19)$$

式中, $w_i = \sum_{u=1}^m \frac{\mathbf{q}_u}{\sqrt{\mathbf{p}_u(\mathbf{y}_0)}} \sigma[b(\mathbf{x}_i) - u]$, $g(x) = -k'(x)$ 。

2.2 目标最佳尺度估计

在文献[10]提出的尺度和方向自适应 Mean Shift 算法基础上提出了一种快速估计出目标最佳尺度的方法,该方法联合候选目标区域的相似性权重分布图的一阶统计特征和候选目标直方图模型与目标直方图模型的相似度精确估计出目标的尺度。

绝大多数视频序列中目标的变化是渐变的,因此,目标的尺度和位置变化可以看做是一个平稳的过程。假设在上一帧中已经得到目标所在的区域,当前帧中让候选目标区域的尺度明显大于上一帧目标区域的尺度。由于相邻帧目标尺度的微小缩放及目标位置的缓慢移动,无论目标尺度变大还是变小,目标将仍然处于该候选目标区域。在 Mean Shift 算法中,初始帧中已经得到目标的颜色直方图模型,在候选目标区域可计算得到区域内像素点颜色值与目标颜色直方图模型对应分量的相似性权值,由此可以得到候选区域的权重分布图。权重分布图描述了候选目标区域内像素点与目标的相似性分布。对于候选目标区域内点 $\mathbf{x}_i$,其权值

$$w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(y_0)}} \sigma[b(x_i) - u] \quad (20)$$

如图 1 所示,图 1(a)是由 4 个颜色灰度组成的目标;图 1(b)是能够把目标完全包含在内的候选区域。图 1(c)(e)(g)分别表示目标不同尺度下的候选目标区域;图 1(d)(f)(h)分别表示利用 Mean Shift 算法原理得到的候选目标区域像素权重分布图。

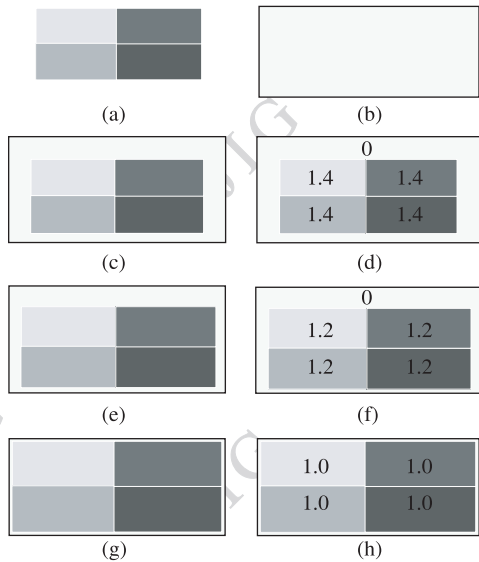


图 1 目标不同尺度下的候选目标区域权重分布图

Fig. 1 The weight image of target candidate region with different scale of target

从图 1 看出,权重分布图权值随目标尺度变化而变化,当候选区域明显大于目标尺度时,候选区域包含有较多背景像素,此时背景像素的权值较小,目标像素的权重较大,并且由于核函数带宽 h 的作用,目标像素的权值大于 1;当目标尺度接近于候选目标区域时,由于没有背景的干扰,目标像素权值近似等于 1。因此,权重分布图可以看做目标在候选区域的密度分布,权重分布图密度值越大的像素点代表该像素点属于目标的概率越大。基于权重分布图,用权重分布图的一阶统计特征来反映目标尺度的大小,即

$$M_0 = \sum_{i=1}^n w(x_i) \quad (21)$$

另一方面,候选目标直方图模型与目标直方图模型的相似度用 $\rho(y)$ 来描述; $\rho(y)$ 越接近于 1 表示两者越相似; $\rho(y)$ 越接近于 0 表示候选目标区域可能包含较多的背景,而候选区域尺度大于目标尺度,

则必有候选区域内的目标尺度较小。由前面分析可知,此时用 M_0 来估计目标尺度会大于其真实尺度。因此, $\rho(y)$ 可以很好地反映用 M_0 来估计目标尺度的可靠性。随着相似度系数 $\rho(y)$ 的增加。 M_0 对目标尺度的估计准确性也增加。为精确估计目标的真实尺度,用 $\rho(y)$ 的函数来修正 M_0 对目标尺度的估计,即

$$A = c(\rho) M_0 \quad (22)$$

式中, $c(\rho)$ 是关于 $\rho(0 \leq \rho \leq 1)$ 的单调递增函数。

$$c(\rho) = \exp\left(\frac{\rho - 1}{\sigma_1}\right), \sigma_1 \text{ 为常数} \quad (23)$$

并且当 ρ 等于 1 时, $c(\rho)$ 取最大值 1; A 是目标尺度的估计值。

当候选区域尺度大于目标尺度时, $c(\rho)$ 小于 1,而 M_0 大于目标的真实尺度,于是可以用 $c(\rho)$ 修正 M_0 对目标尺度的估计。当候选区域尺度等于目标尺度时, $c(\rho)$ 等于 1, M_0 即为目标的尺度值,实验证明尺度估计因子 σ_1 值取 1 到 2 之间时,该方法对目标的尺度估计效果较好^[10]。

3 本文跟踪算法

3.1 基本原理

3.1.1 确定目标位置

本文目标状态考虑目标的位置和尺度。目标状态用一个矩形框表示 $R = \{l, t, w, h\}$, 其中 $\{l, t\}$ 是矩形框中心位置坐标, $\{w, h\}$ 是矩形框的长度和宽度,约定目标长宽比为常数 $K = h/w$ 。设定空间位置变化范围为目标大小的 2 倍,尺度缩放范围为原来的 0.8 ~ 1.2 倍。假设已经获得目标在当前帧的状态 $R = \{l, t, w, h\}$ 及目标直方图模型 q_u 。定义训练样本为 $X \in \mathbf{R}^{2w \times 2h}$ 。每个样本对应的矩形框为

$$R_{ij} = \{l - w/2 + i, t - h/2 + j, w, h\} \quad (24)$$

式中, $0 \leq i \leq w, 0 \leq j \leq h$ 。

训练样本表示为 $X_{ij} = I(R_{ij})$

对应的响应 $y \in \mathbf{R}^{2w \times 2h}$ 表示为

$$y_{ij} = \exp\left(-\frac{(i-w)^2 + (j-h)^2}{2\delta_p^2}\right) \quad (25)$$

式中, δ_p 控制空间位置对响应的强度。图 2 为在固定尺度上的训练样本。

根据第 2 节介绍的回归方法,利用训练数据

$\{x, y\}$, 得到核岭回归的解如式(5), 定义变量

$$\alpha = (\lambda I_N + K)^{-1} y \quad (26)$$

模型解可重写为

$$\hat{w} = x^T \alpha = \sum_{i=1}^N \alpha_i x_i \quad (27)$$

对于在下一帧图像中新的候选样本 x' , 根据模型可以预测该样本属于目标的响应, 即

$$\hat{y} = \hat{w}^T x' = \sum_{i=1}^N \alpha_i x_i^T x' = \sum_{i=1}^N \alpha_i \kappa(x^i, x') \quad (28)$$

根据卷积定理得到对偶变量的频域求解, 即

$$\alpha = (\lambda I_N + K)^{-1} y = F^{-1} \left(\frac{F(y)}{F(K) + \lambda} \right) \quad (29)$$

预测值为

$$\hat{y} = F^{-1}(F(K)F(\alpha)) \quad (30)$$

响应最大的位置 $\hat{y} = (l', t')$ 即目标当前中心位置。

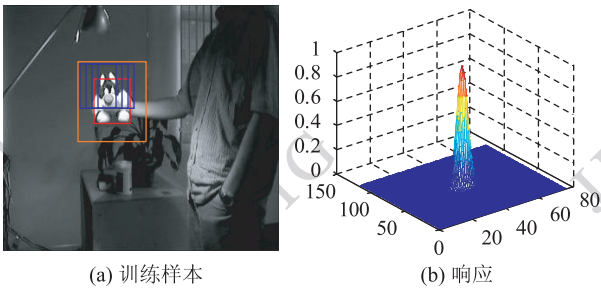


图2 训练样本示例

Fig. 2 Training sample example

3.1.2 确定目标尺度

基于目标的中心位置 $\hat{y}$, 利用式(16)(17)分别求得以 $\hat{y}$ 为中心的候选目标区域的直方图模型 $p_u(\hat{y})$ 及其与目标直方图模型 q_u 的相似度 ρ , 进而可用式(20)计算候选目标区域的权重分布图。最后利用式(22)对目标的尺度进行精确估计。因为本文目标状态用矩形框表示及目标长宽比限定为恒定值 K , 则目标长宽分别为

$$h' = \sqrt{KA}, w' = \sqrt{A/K} \quad (31)$$

于是, 目标的当前状态为 $R = \{l', t', w', h'\}$ 。

3.1.3 模型更新

当前帧跟踪成功后, 为了适应目标变化及减少背景的干扰, 设置了一个模型更新因子 β 进行模型 $\hat{w}$ 实时更新, 即

$$\hat{w}' = (1 - \beta) \hat{w}^{t-1} + \beta \hat{w}^t \quad (32)$$

3.2 算法步骤

本文算法完整描述如下:

1) 标定目标的初始状态获得训练数据并利用式(27)(29)计算回归跟踪模型, 根据式(15)计算目标的颜色直方图模型 q_u ;

2) 根据式(30)搜索目标, 确定目标中心位置 $\hat{y}$;

3) 根据式(16)(17)计算候选目标颜色直方图模型 $p_u(\hat{y})$ 及其与目标颜色直方图模型 q_u 的相似度 ρ ;

4) 根据式(20)计算候选目标区域权重分布图;

5) 根据式(22)精确估计目标的尺度值 A , 由式(31)计算目标的长和宽, 确定目标的跟踪结果;

6) 重新获得训练数据, 计算当前帧回归跟踪模型, 根据式(32)进行模型更新, 继续下一帧跟踪。

本文算法流程如图3所示。

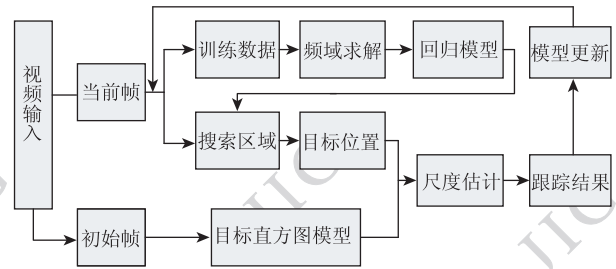


图3 本文算法的流程框图

Fig. 3 Flow chart of the our tracking algorithm

4 实验结果及分析

本文实验所使用的测试视频均是包含尺度变化的常用视频, 实验中设置空间位置变化范围为目标大小的2倍, 尺度变化范围为原来的0.8~1.2倍, 通过大量实验表明正则化参数为0.01时跟踪效果最好, 高斯核方差 σ 、尺度估计因子 σ_1 ($1 \leq \sigma_1 \leq 2$) 和模型更新因子 β 在不同视频中具有不同的最佳值。实验中其他对比算法均来自相应作者的公开代码。所有实验都在联想 CPU-E5300 @ 2.60 GHz、3.25 GB 内存的台式机进行, 算法通过 MATLAB2009a 实现。

4.1 定性分析

本文采用的对比算法分别是增量试探法 (incre)^[3]、CBWH^[8]、IVT^[12]、DFT^[7]、DSST^[5]; 其中增量试探法是一种对目标尺度变化具有适应能力的传统跟踪算法, CBWH、IVT、DFT 及 DSST 是目前适用性较广的主流跟踪算法。图4是跟踪实验的部分结

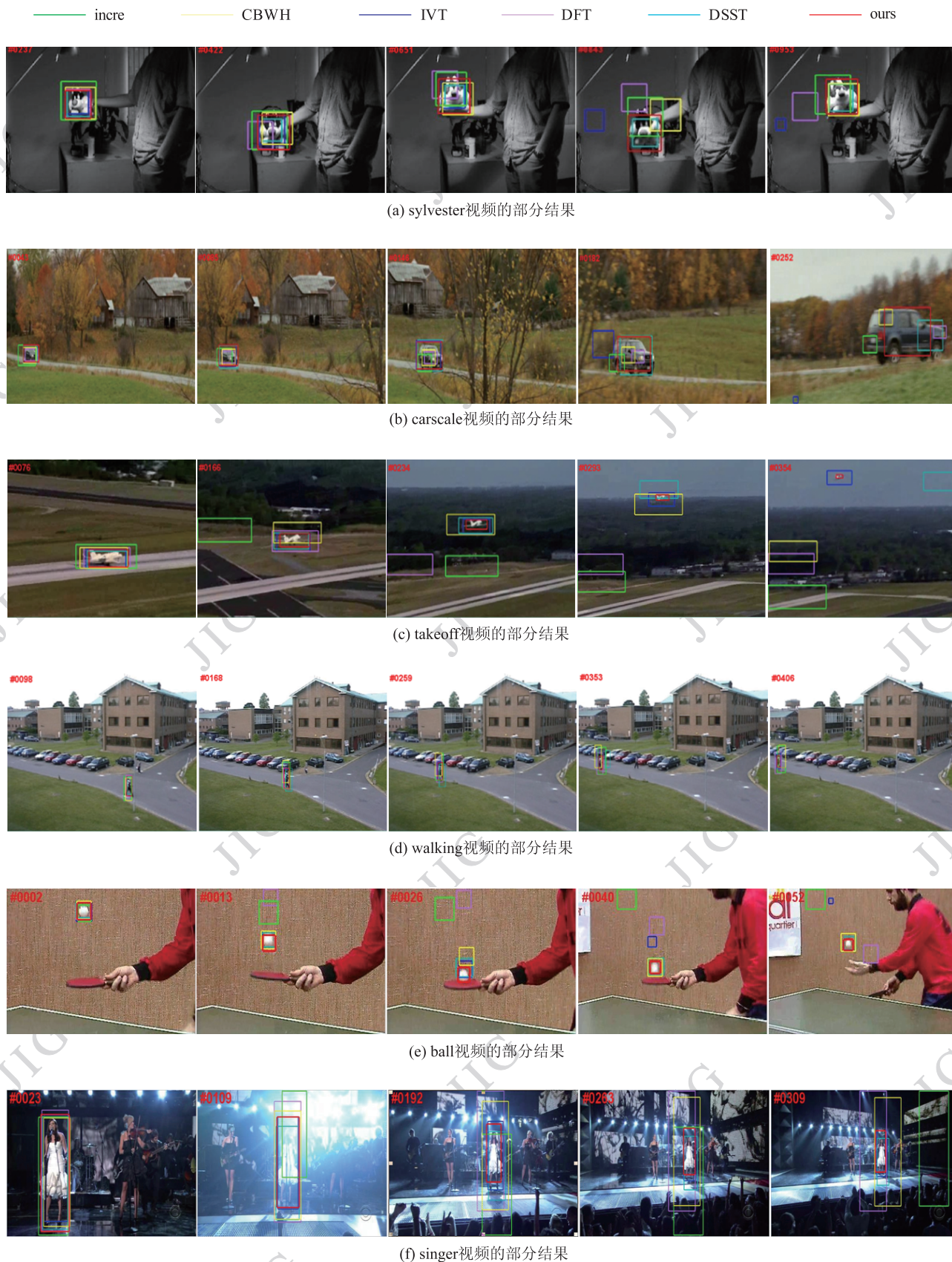


图4 跟踪算法性能的定性比较

Fig. 4 Qualitative comparison of the four tracking algorithms

果,不同算法的跟踪结果分别采用不同颜色的实线矩形框表示,左上角是图像帧数。

图 4(a)是 *sylvester* 视频的部分跟踪结果,由于后部分视频中目标由正面转向侧面,姿态改变太大,故本文仅截取视频的前 960 帧进行跟踪实验。从实验结果可以看出,CBWH 和 DFT 由于不能适应目标的尺度变化,容易跟踪失败;增量试探法在目标姿态及光照发生较大变化时跟踪误差较大;在第 843 帧中显示,当目标姿态发生较大变化时,IVT 跟踪失败;DSST 算法对姿态及光照变化具有一定的适应能力,但不能很好地适应目标尺度变化;而本文算法能很好地适应目标尺度变化,具有较好的跟踪精度。图 4(b)是 *carscale* 视频跟踪部分结果,视频中的目标存在遮挡和姿态及尺度变化,从 182 帧中可以看出,当目标被严重遮挡时,IVT 重构目标的误差变大,跟踪失败;第 252 帧结果显示目标姿态由前侧面变为正侧面,即目标姿态发生较大变化时,增量试探法、CBWH、DFT、DSST 都出现了跟踪偏差较大的情况,而本文算法对目标的变化适应性最好。图 4(c)是 *takeoff* 视频的部分跟踪结果,从结果可以看出,前部分视频中目标姿态及背景发生较小变化时,几种算法都能成功跟踪目标,但是在后部分视频中当背景发生较大变化时,增量试探法、CBWH、DFT、DSST 都跟踪失败,IVT 和本文算法都能较好地跟踪目标,但本文算法对目标尺度的明显变化适应能力更好,因此跟踪效果更好。图 4(d)是 *walking* 视频的部分跟踪结果,视频中目标存在尺度及背景的干扰,从第 168 帧可以看出,当目标背景出现与目标相似性较高的外物干扰时,增量试探法、CBWH、DSST 都出现了较大的跟踪偏差。IVT、DFT 及本文算法都能成功跟踪目标,但本文算法对目标的尺度具有更好的适应能力,因此具有更好的跟踪效果。图 4(e)是 *ball* 视频的部分跟踪结果,视频中目标运动速度时快时慢,当球上升时速度不断减小,下降时速度不断增大,同时目标存在尺度和运动模糊,可以看出,当目标运动较快导致模糊时,增量试探法、IVT、DFT 均丢失目标,如第 40 帧所示,CBWH、DSST 和本文算法都能较好跟踪目标,但本文算法能较好适应尺度变化,跟踪效果更好。图 4(f)是 *singer* 视频部分跟踪结果,视频存在尺度变化及复杂背景干扰,从第 192 帧可以看出,背景干扰严重时,增量试探法、DFT 跟踪偏差较大,CBWH 由于不能适应目标

尺度变化,跟踪偏差较大。IVT、DSST 和本文算法都能成功跟踪目标。

4.2 定量分析

采用 3 个指标来衡量本文算法与对比算法的优劣:平均每帧运算时间、中心位置误差及覆盖率。其中,平均每帧运算时间衡量跟踪效率,中心位置误差和覆盖率衡量跟踪精度。

4.2.1 跟踪效率

平均每帧运算时间表示跟踪算法进行跟踪每帧所需时间,值越小跟踪效率越高。表 1 所示为跟踪算法平均每帧运算时间。可以看出本文算法与对比算法相比跟踪实时性具有明显优势。

表 1 平均每帧运行时间

Table 1 Average running time of per frame

/(s/帧)						
视频	incre	CBWH	IVT	DFT	DSST	本文
<i>sylvester</i>	0.187	0.309	0.129	0.552	1.718	0.033
<i>carscale</i>	0.128	0.193	0.195	0.501	1.874	0.060
<i>takeoff</i>	0.293	0.567	0.154	0.783	1.894	0.096
<i>walking</i>	0.174	0.318	0.177	0.626	1.954	0.068
<i>ball</i>	0.212	0.166	0.126	0.467	1.622	0.026
<i>singer</i>	0.196	0.402	0.157	0.622	1.783	0.081

注:每个视频对应最优算法标为黑体,次优算法标为斜体。

4.2.2 跟踪精度

中心位置误差表示目标中心位置与真实目标中心位置之间的误差,其值越小跟踪精度越高;覆盖率表示跟踪结果与真实目标的重叠度,值越大跟踪精度越高。表 2 和表 3 分别表示平均中心位置误差值及平均覆盖率;图 5 和图 6 分别表示中心位置误差曲线及覆盖率曲线。

表 2 平均中心位置误差值

Table 2 Average center location error

名称	incre	CBWH	IVT	DFT	DSST	本文
<i>sylvester</i>	13.901	11.372	32.818	29.841	4.136	4.872
<i>carscale</i>	38.809	26.337	70.825	23.239	15.192	11.079
<i>takeoff</i>	119.331	35.597	7.260	98.189	38.157	5.072
<i>walking</i>	8.268	14.488	2.519	2.001	7.429	1.925
<i>ball</i>	53.426	2.440	18.681	33.786	4.769	2.226
<i>singer</i>	65.870	12.765	8.002	18.882	14.023	10.482

注:每个视频对应最优算法标为黑体,次优算法标为斜体。

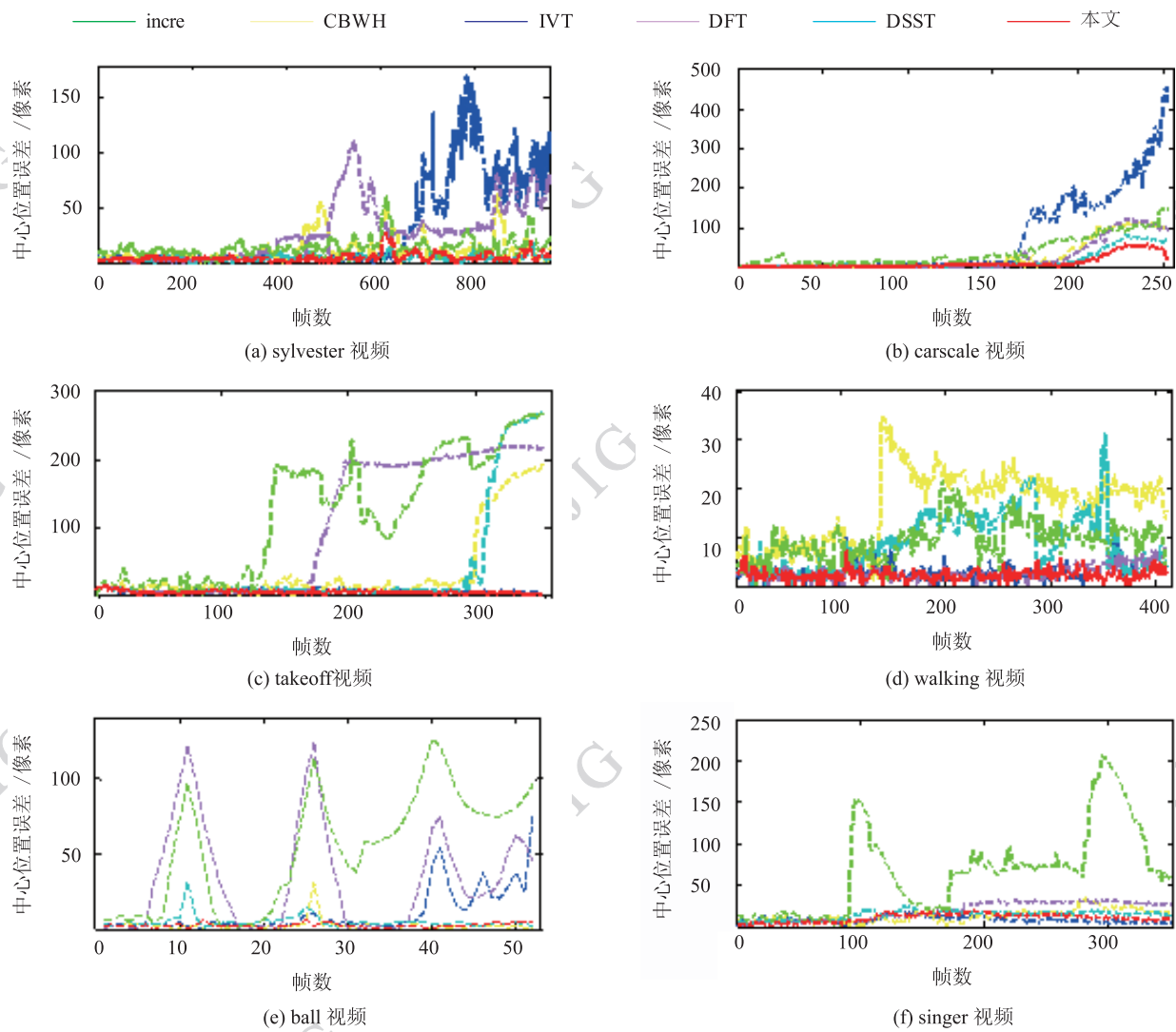


图5 中心位置误差曲线
Fig. 5 Center location error curves

表3 平均覆盖率
Table 3 Average overlap

名称	incre	CBWH	IVT	DFT	DSST	本文
sylvester	0.563	0.641	0.521	0.393	0.799	0.801
carscale	0.297	0.598	0.525	0.700	0.773	0.789
takeoff	0.198	0.401	0.606	0.376	0.515	0.690
walking	0.559	0.414	0.839	0.816	0.517	0.850
ball	0.203	0.819	0.557	0.372	0.693	0.823
singer	0.462	0.749	0.802	0.642	0.671	0.804

注:每个视频对应最优算法标为黑体,次优算法标为斜体。

图5 表示跟踪结果的中心位置误差曲线,其值

越小表明跟踪位置与真实位置越接近,跟踪精度越高;本文算法相比其他对比算法中心位置误差始终保持较低水平,表2所示为平均中心位置误差值, Sylvester 视频中,本文算法平均中心位置误差值略高于 DSST 算法, singer 视频中,本文算法平均中心位置误差值略高于 IVT 算法,其他视频中本文算法均是最优算法。

图6 表示算法的跟踪覆盖率曲线,值越大表示跟踪结果与真实目标重叠度越大,跟踪精度越好;从曲线可以看出本文算法相比其他对比算法始终保持较好的覆盖率。从表3所示的平均覆盖率可看出本文算法在所有视频中始终具有最高的平均覆盖率。

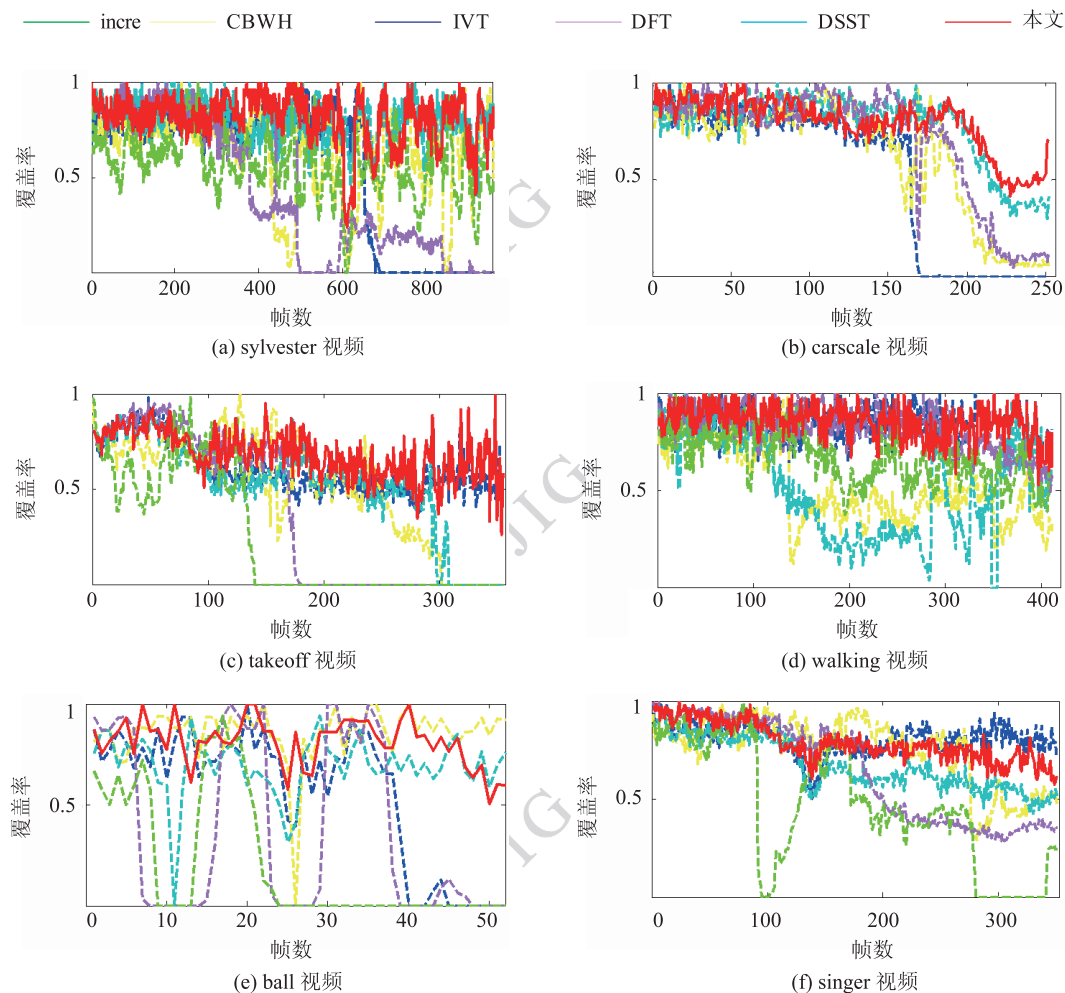


图6 覆盖率曲线

Fig. 6 Overlap curves

5 结论

针对传统跟踪算法不能很好适应目标尺度变化的问题,提出了一种基于快速傅里叶变换的尺度自适应回归跟踪算法。该算法在建立目标跟踪回归模型的基础上,提出利用候选目标区域像素点的权重分布图有效解决了目标尺度变化的问题。

实验结果表明,本文算法不仅能较好适应目标尺度及背景变化,而且通过适应目标尺度的变化,提高了跟踪的鲁棒性;同时,利用快速傅里叶变换将时域运算变换到频域运算,提高了跟踪效率,具有很好的实时性,说明该算法在实时视频跟踪中具有很好的应用前景。

在仿真中发现,当目标具有较大角度旋转时,本

文算法跟踪效果不是很好。因此,如何适应目标的旋转变化问题将是下一步的工作。

参考文献 (References)

- [1] Yang H X, Shao L, Zheng F, et al. Recent advances and trends in visual tracking: a review [J]. Neurocomputing, 2011, 74: 3823-3831.
- [2] Hou Z Q, Han C Z. A survey of visual tracking [J]. Acta Automatica sinica, 2006, 32(4): 603-617. [侯志强, 韩崇昭. 视觉跟踪技术综述 [J]. 自动化学报, 2006, 32(4): 603-617.]
- [3] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.
- [4] Collins R T. Mean-Shift blob tracking through scale space [C]// Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, Madison, WI, United States: IEEE, 2003: 234-240.

- [5] Zhuang B H, Lu H C, Xiao Z Y, et al. Visual tracking via discriminative sparse similarity map [J]. IEEE Transactions on Image Processing, 2014, 23(4):1872-1881.
- [6] Wang D, Lu H C, Yang M. Online object tracking with sparse prototypes [J]. IEEE Transactions on Image Process, 2013, 22(1):314-325.
- [7] Sevilla L, Learned E. distribution fields for Tracking. [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, United states: IEEE, 2012: 1910-1917.
- [8] Ning J F, Zhang L, Zhang D, et al. Robust mean shift tracking with corrected background-weighted histogram. [J]. IET Computer Vision, 2012, 6(1): 62-69.
- [9] Henriques J F, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels [C]// Proceedings of European Conference on Computer Vision. Berlin: Springer Verlag, 2012: 702-715.
- [10] Ning J F, Zhang L, Zhang D, et al. Scale and orientation adaptive mean shift tracking [J]. IET Computer Vision, 2012, 6(1):52-61.
- [11] Li Q, Shao C. Nonlinear systems identification based on kernel ridge regression and its application [J]. Journal of System Simulation, 2009, 21(8): 2152-2155. [李琦, 邵诚. 基于核岭回归的非线性系统辨识及其应用 [J]. 系统仿真学报, 2009, 21(8): 2152-2155.]
- [12] Ross D, Lim J, Lin R S, et al. Incremental learning for robust visual tracking [J]. International Journal of Computer Vision, 2008, 77(1-3): 125-141.