

# 视觉量子目标跟踪方法\*

姜文涛<sup>1</sup>, 刘万军<sup>1</sup>, 袁 姮<sup>2</sup>, 张海涛<sup>1</sup>



<sup>1</sup>(辽宁工程技术大学 软件学院, 辽宁 葫芦岛 125105)

<sup>2</sup>(辽宁工程技术大学 研究生院, 辽宁 葫芦岛 125105)

通讯作者: 姜文涛, E-mail: lntuwulue@163.com, http://www.lntu.edu.cn

**摘要:** 为了解决变结构目标跟踪过程中目标失跟率较高的问题,提出了一种基于视觉量子(vision quantum,简称VQ)的目标跟踪方法.该方法首先在图像内自上而下地辐射视觉量子采集灰度信息,统计量子内部概率密度较大的灰度级和分布区域;然后计算视觉量子的量子频率,归一化量子频率系数,滤除系统噪音和杂波干扰,利用频率阶跃不变性移动视觉量子至平衡状态,将达到量子平衡状态的视觉量子组成量子簇;最后,以该量子簇作为候选目标信息,采用极大似然估计预测运动目标状态,以预测结果作为下一帧图像中视觉量子移动的参考值,并进一步验证移动后的视觉量子是否达到量子平衡状态,以确保目标跟踪有效性.该方法抓住了变结构运动目标前景与背景交界处具有量子频率阶跃不变性的特点,继而将阶跃不变特征采用具有独立性和约束性的视觉量子进行描述,可以有效地消除形状变化、尺度变化等变结构因素对运动目标跟踪的影响,失跟率较低.同时,由于视觉量子数据量较小,计算复杂度较低,其跟踪实时性较高.大量实验测试结果表明,该方法对变结构目标跟踪具有很好的适应性、实时性和鲁棒性.

**关键词:** 视觉量子;频率阶跃;量子平衡;量子簇;目标跟踪

中图法分类号: TP391

中文引用格式: 姜文涛,刘万军,袁姮,张海涛.视觉量子目标跟踪方法.软件学报,2016,27(11):2961-2984. http://www.jos.org.cn/1000-9825/4931.htm

英文引用格式: Jiang WT, Liu WJ, Yuan H, Zhang HT. Object tracking method based on vision quantum. Ruan Jian Xue Bao/ Journal of Software, 2016, 27(11): 2961-2984 (in Chinese). http://www.jos.org.cn/1000-9825/4931.htm

## Object Tracking Method Based on Vision Quantum

JIANG Wen-Tao<sup>1</sup>, LIU Wan-Jun<sup>1</sup>, YUAN Heng<sup>2</sup>, ZHANG Hai-Tao<sup>1</sup>

<sup>1</sup>(School Of Software, Liaoning Technical University, Huludao 125105, China)

<sup>2</sup>(Graduate school, Liaoning Technical University, Huludao 125105, China)

**Abstract:** An approach to object tracking based on vision quantum is proposed in this paper in order to solve the high loss-tracking rate in variable structure object tracking. First, the gray information is detected in an image from top to bottom with vision quantum, and the distribution area and gray levels of larger probability density are counted in the vision quantum. Then all the energy frequencies of the visual quantum are calculated such that the weaker energy frequency gradient is removed by filtration and the stronger frequency gradient of vision quantum that the distribution of high frequency information account for half quantum area is reserved. The quantum cluster is composed of vision quantum with the same frequency variation. Finally, taking quantum cluster as candidate object information, the state

\* 基金项目: 国家自然科学基金(61172144); 国家高技术研究发展计划(863)(13-2025); 辽宁省教育厅科学研究项目(LJYL049); 辽宁省科技攻关计划(2012216026)

Foundation item: National Natural Science Foundation of China (61172144); National High-Tech R&D Program of China (863) (13-2025); Science Research Project of Liaoning Provincial Department of Education (LJYL049); Science and Technology Foundation of Liaoning Province (2012216026)

收稿时间: 2014-04-27; 修改时间: 2015-06-24, 2015-06-30; 采用时间: 2015-10-09; jos 在线出版时间: 2015-11-15

CNKI 网络优先出版: 2015-11-16 09:22:23, http://www.cnki.net/kcms/detail/11.2560.TP.20151116.0922.004.html

of moving object is predicted with maximum likelihood estimation and the forecast results are served as moving reference position of vision quantum in the next frame. Further verification of the visual quantum balance state is made to ensure the effectiveness of object tracking. This method catches the point that the variable structure moving object has the feature of the energy frequency step invariance at the juncture pixels of the foreground and background. It can effectively overcome the changes in shape, scale and other factors that influence the moving object tracking, achieving lower loss-tracking rate and lower computational complexity by using independent and continuous visual quantum to describe the step invariant feature. Experimental results show that the proposed approach has good adaptability to variable structure tracking with real-time and robust tracking performance.

**Key words:** vision quantum; frequency step; quantum balance; quantum cluster; object tracking

运动目标跟踪 MOT(moving object tracking)作为目前国内外计算机视觉领域的研究热点之一<sup>[1,2]</sup>,经过 50 余年的研究发展,已经取得了较大的成绩.它是一种利用计算机与摄像机模仿“生物视觉系统”对感兴趣目标进行分析、识别、跟踪及测量的技术,具有无接触、隐蔽性强、精准性高的技术优点<sup>[3-7]</sup>.目标跟踪应用广泛,在仿生机器人、无人驾驶汽车、智能视频监控等自动感知领域具有重要的研究和应用价值.

目前,目标跟踪技术的研究主要有两大方向:一种是建立在运动信息分析基础上的跟踪方法,另一种是基于模型分析的目标跟踪方法.

利用运动信息分析的目标跟踪方法研究较多,已形成了众多的算法和标准.例如,Huttenlocher 等人首先利用 Hausdorff 距离的定义与计算方法进行序列图像间的去相关,然后对残余误差图像采用空间去相关实现了物体的跟踪算法<sup>[8,9]</sup>;在 Huttenlocher 的基础上,Haritaoglu 等人通过建立双峰分布的统计背景模型来实现目标检测,并为目标建立外观模型,取得了更好的效果<sup>[10]</sup>;在进行空间域上去相关的基础上,Stauffer 等人提出了自适应混合高斯模型,建立运动目标与前景目标间的对应关系来实现对目标的跟踪<sup>[11]</sup>;在自适应模型的基础上,Rezaeian 等人利用观测模型联合收敛(convergence observation model of joint,简称 COMJ)技术对目标进行检测,然后应用序贯蒙特卡罗方法(sequential Monte Carlo,简称 SMC)对目标进行拓展跟踪,获得了较好的效果<sup>[12]</sup>.目标跟踪方法也不仅仅限于空间域上去相关方法,国内的闫小喜等人提出了基于杂波强度在线估计的目标跟踪方法<sup>[13]</sup>;Juan 等人提出了基于视频对象运动的模糊推理方法,取得了较好的跟踪效果,但存在目标分割完整性欠缺的问题<sup>[14]</sup>;Airouche 等人对此进行了改进,采用运动区域增长和先验知识结合的方法使视频对象的提取较为精确,并具有较好的鲁棒性能<sup>[15]</sup>.基于运动分析的目标跟踪方法原理简单,实现较为方便,但是对于相似运动目标难以进行区分,容易粘连和失跟.

另一种基于模型的目标跟踪方法可以克服相似运动信息的干扰.近年来,基于模型的目标域密度估计方法(mean shift)成为研究的热点<sup>[16-21]</sup>.目标域密度估计是由 Fukunaga 和 Hostetler 在 1975 年提出的一种无监督聚类方法,它使每一个点“漂移”到密度函数的局部极大值点.Comanicu 等人首先将 mean shift 算法应用到目标跟踪领域中来,他们利用 Bhattacharyya 系数作为目标模型域候选目标之间的相似性度量标准,用 mean shift 算法来搜索最优候选目标,获得了很好的跟踪效果<sup>[22,23]</sup>,但不足之处是计算量较大,影响了跟踪的准确性和实时性.为了弥补特征信息不足造成的偏差,Yang 等人提出了空域特征中的相似性度量方法,允许 mean shift 跟踪更一般的运动目标<sup>[24]</sup>,但是对于相似目标的衔接不能进行有效判别.在此基础上,Jeong 等人提出了一个鲁棒的、实时的高斯圆柱模型,通过 mean shift 实现目标跟踪,以解决不规则光照变化下目标衔接带来的影响<sup>[25]</sup>,但是没有解决形状变化导致的误跟问题.高斯模型的不足是基函数需要计算与原始数据相关的协方差矩阵和特征矢量,这些计算在应用中通常是较难实现的.因此,人们往往利用目标参数分布模型来代替高斯模型变换的去相关.例如,Davy 等人通过在高斯混合模型中引入先验参数的分布模型,采用混合监督对目标进行分类,可以同时跟踪多种运动目标,对于一定角度内的外观形变具有鲁棒性<sup>[26]</sup>,但稳定性较低,容易漂移;为了提高目标检测与跟踪的稳定性,Liam 等人提出线性回归和自适应快速同步建模方法,通过自适应运动目标的外观模型进行目标更新,增强目标跟踪的稳定性<sup>[27]</sup>;Ying 等人提出对背景变化的运动目标前景信息进行混合高斯处理<sup>[28]</sup>,将多种纹理直方图和颜色直方图建立目标运动模型,选择近似模型进行视频目标跟踪,取得了较好的跟踪效果.另外,不同的图像特征描述方法适合于不同的场合,因此,根据场景来自适应选择特征描述方法是非常重要的研究内容,例如,Yi

等人研究了目标跟踪中适合于场景变化的在线特征选择机制,以提高跟踪性能<sup>[29]</sup>。其次,在目标跟踪中引入随机有限集的目标模型是一种行之有效的方法,对目标模型进行学习后,确定相应特征和参数,可提高对周围杂波环境的适应性。例如,吕学斌等人通过引入概率密度滤波器,建立判决门限的跟踪策略,以此提高对周围杂波环境的适应性<sup>[30]</sup>。基于模型更新的目标跟踪方法是计算机视觉领域内的主流方法,例如 Kalal 等人的 TLD(tracking-learning-detection)方法<sup>[31]</sup>、Rui 等人的 OLSL(online latent structural learning)方法<sup>[32]</sup>、Junseok 等人的 WLMC(Wang-Landau Monte Carlo)方法<sup>[33]</sup>,对相似运动信息和低速变结构运动具有较强的抗干扰性,可以较好地跟踪运动目标,但是在目标发生快速变结构运动,例如目标快速旋转导致目标形状和尺度变化幅度较大时,却不能进行长时间的有效跟踪,容易发生漂移和误跟。

与生物的视觉系统相比,现有目标跟踪方法的普遍弱点是缺乏与感兴趣目标及其周围环境同步调节的能力,而生物的视觉系统可以自然地感知目标和环境的变化,做到协同调节,以形成对感兴趣目标的最大关注度。蝇眼是一种复眼结构,由多个六边形结构的小眼睛拼成,每个小眼睛都可以独立成像,协同工作,最后形成一个影像,而人眼内部也是由多个分布密集的感光神经单元组成,由多个感光单元协同完成目标成像,生物“视觉单元”如图 1 所示。这种生物视觉系统可以同时抓住目标多个局部边缘信息的变化,因此对目标整体形状和尺度的变化感知十分灵敏,对其变化的幅度也具有很好的鲁棒性<sup>[34,35]</sup>。

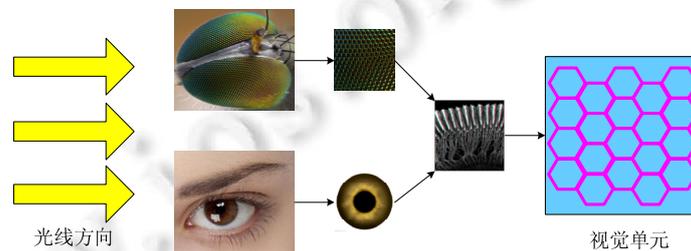


Fig.1 Biological “visual unit” photosensitive diagram

图 1 生物“视觉单元”感光示意图

在生物视觉系统的启发下,张讲社等人<sup>[36,37]</sup>模拟生物复眼结构,提出一种基于方向量子化表示理论的并行数据处理方法,利用最小矢量集合模拟生物“视觉单元”,在机器人运动控制和自适应图像处理领域取得了较好的应用效果。在计算机视觉领域中,如果现有的目标跟踪方法具有生物视觉系统的这种多个“视觉单元”的协同工作能力,将对目标跟踪效果的提升起到不可限量的作用。基于此,本文在生物“视觉单元”协同工作的方向上,提出了一种视觉量子(vision quantum,简称 VQ)的概念,并应用到目标跟踪中。

视觉量子目标跟踪方法是建立在目标子空间协同分析的整体框架上,利用辐射视觉量子采集灰度信息计算能量频率,保留达到量子平衡的视觉量子,将频率变化幅度相同的视觉量子组成量子簇,以该量子簇作为候选目标信息,采用极大似然估计预测运动目标状态,重复验证移动后的视觉量子是否达到量子平衡状态,确保目标跟踪有效性。与现有目标跟踪方法相比,该方法的优势在于:

- (1) 抓住了变结构运动目标前景与背景交界处具有的能量频率阶跃不变性的特点;
- (2) 将阶跃不变特征采用多个具有独立性和约束性的视觉量子进行描述,可以有效消除形状变化、尺度变化等变结构因素对运动目标跟踪的影响,失跟率较低;
- (3) 视觉量子数据量较小,同时不需要存储目标多种姿态模型,计算复杂度和空间复杂度较低,跟踪实时性较高。

视觉量子对非连续的刚性突变运动不具有跟踪意义,作为视觉量子实现长时间稳定跟踪的基础,目标运动轨迹及其频域变换必须具有连续的光滑特性,才能确保对频率阶跃计算的有效性,达到量子反复平衡,实现对目标的稳定跟踪。因此,本文首先给出了目标运动可微分定理及其证明过程;在此基础上,提出了视觉量子理论模型及其相关概念的定义;然后阐述了该方法的基本实现流程,并给出了变结构运动目标视觉量子跟踪效果;最

后,在多个变结构目标运动视频数据库中对本文 VQ 方法、TLD 方法<sup>[31]</sup>、OLSL 方法<sup>[32]</sup>、WLMC 方法<sup>[33]</sup>进行了跟踪实验,并给出实验结果和对比分析.

### 1 目标运动可微分定理

目标运动轨迹及其频域变换的连续光滑特性如图 2 所示.

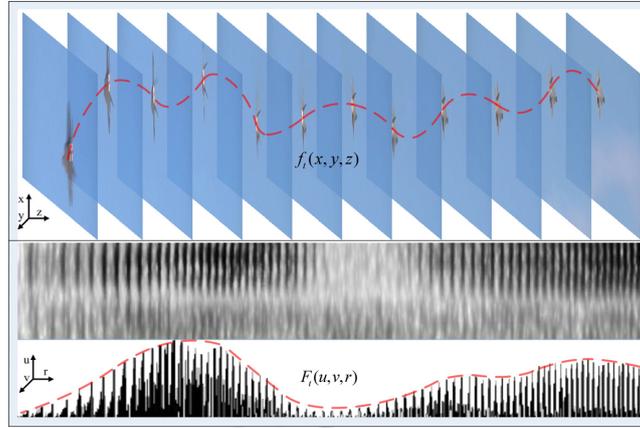


Fig.2 Continuous smooth sketch map of object trajectory and its frequency domain transformation

图 2 目标运动轨迹及其频域变换的连续光滑特性示意图

**定理 1.** 在目标运动过程中,其运动轨迹  $\varphi$  在时域内是连续光滑的,轨迹记为  $\varphi=f_i(x,y,z)$ ,则轨迹  $\varphi$  的偏导数  $\partial\varphi/\partial x, \partial\varphi/\partial y, \partial\varphi/\partial z$  在其定义域内是连续的,即轨迹  $\varphi$  是可微分的.

证明:设点  $(x,y,z)$  为自变量定义域内任意一点,则  $(x+\Delta x, y+\Delta y, z+\Delta z)$  为该点邻域内任意一点,由数学理论可知,轨迹全增量  $\Delta\varphi$  为

$$\begin{aligned} \Delta\varphi &= f_i(x+\Delta x, y+\Delta y, z+\Delta z) - f_i(x, y, z) \\ &= [f_i(x+\Delta x, y+\Delta y, z+\Delta z) - f_i(x, y+\Delta y, z+\Delta z)] + [f_i(x, y+\Delta y, z+\Delta z) - f_i(x, y, z+\Delta z)] + \\ &\quad [f_i(x, y, z+\Delta z) - f_i(x, y, z)]. \end{aligned}$$

当  $\Delta x \rightarrow 0, \Delta y \rightarrow 0, \Delta z \rightarrow 0$  时,轨迹全增量  $\Delta\varphi$  可以表示为

$$\Delta\varphi = f'_x(x, y, z)\Delta x + f'_y(x, y, z)\Delta y + f'_z(x, y, z)\Delta z + \sigma_1\Delta x + \sigma_2\Delta y + \sigma_3\Delta z,$$

其中,  $\sigma_1, \sigma_2, \sigma_3$  分别为  $(\Delta x, \Delta y, \Delta z), (\Delta y, \Delta z), (\Delta z)$  的函数,且当  $\Delta x \rightarrow 0, \Delta y \rightarrow 0, \Delta z \rightarrow 0$  时,  $\sigma_1 \rightarrow 0, \sigma_2 \rightarrow 0, \sigma_3 \rightarrow 0$ . 此时,全增量  $\Delta\varphi$  逐渐趋于 0,即,  $\varphi=f_i(x,y,z)$  在任意点  $(x,y,z)$  可微分. □

**定理 2.** 目标运动轨迹  $\varphi$  的傅里叶变换  $F_i(u, v, r) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_i(x, y, z) e^{-2\pi j(ux/1+vy/m+rz/n)} dx dy dz$  在时域内是连续光滑的,则函数  $F_i(u, v, r)$  的偏导数  $\partial F_i/\partial u, \partial F_i/\partial v, \partial F_i/\partial r$  在其定义域内是连续的,即函数  $F_i(u, v, r)$  是可微分的.

证明:设点  $(u, v, r)$  为自变量定义域内任意一点,则  $(u+\Delta u, v+\Delta v, r+\Delta r)$  为该点邻域内任意一点,由数学分析理论可知,函数全增量为

$$\Delta F_i = F'_u(u, v, r)\Delta u + F'_v(u, v, r)\Delta v + F'_r(u, v, r)\Delta r + \phi,$$

其中,

$$\begin{aligned} \phi &= [F'_u(u + \alpha\Delta u, v + \Delta v, r + \Delta r) - F'_u(u, v, r)]\Delta u + [F'_v(u + \Delta u, v + \beta\Delta v, r + \Delta r) - F'_v(u, v, r)]\Delta v + \\ &\quad [F'_r(u + \Delta u, v + \Delta v, r + \gamma\Delta r) - F'_r(u, v, r)]\Delta r. \end{aligned}$$

由柯西不等式:

$$\phi \leq [(F'_u(u + \alpha\Delta u, v + \Delta v, r + \Delta r) - F'_u(u, v, r))^2 + (F'_v(u + \Delta u, v + \beta\Delta v, r + \Delta r) - F'_v(u, v, r))^2 + (F'_r(u + \Delta u, v + \Delta v, r + \gamma\Delta r) - F'_r(u, v, r))^2]^{1/2} \times \sqrt{\Delta u^2 + \Delta v^2 + \Delta r^2}.$$

由于指数函数  $e^{-2\pi j(ux/l+vy/n+rz/n)}$  的偏导是连续的,当  $\Delta u \rightarrow 0, \Delta v \rightarrow 0, \Delta r \rightarrow 0$  时,  $(\Delta u^2 + \Delta v^2 + \Delta r^2) \rightarrow 0$ , 且

$$\left( \frac{\phi}{\sqrt{\Delta u^2 + \Delta v^2 + \Delta r^2}} \right) \rightarrow 0.$$

即  $F_i(u, v, r)$  在任意点  $(u, v, r)$  可微分. □

## 2 视觉量子目标跟踪方法

为了清晰地描述视觉量子目标跟踪方法,将其中涉及到的一些定义和符号给出,见表 1.

**Table 1** Definitions and symbols

表 1 定义和符号

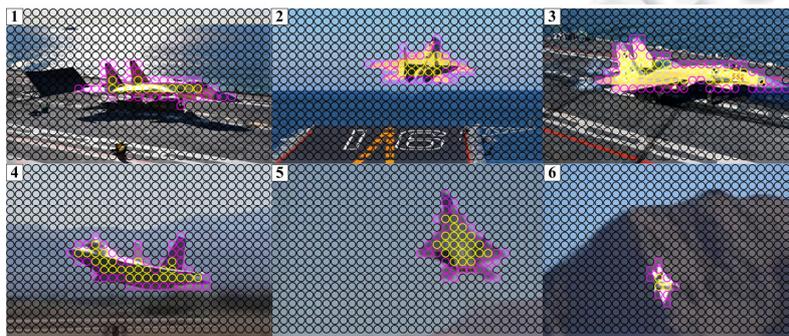
符号	定义	符号	定义	符号	定义
$VFS$	视觉场空间	$G(x, y)$	像素点的灰度	$O'(X', Y')$	边缘区域几何中心
$VQ(V)$	视觉量子(视觉采样单元)	$S$	视觉量子几何面积	$\delta$	移动向量
$EE$	熵能量	$\gamma$	单位灰度能量	$QC$	量子簇
$\zeta$	信息熵	$P$	量子频率	$\alpha_{kl}$	复高斯随机变量
$A$	信息熵极大值的灰度阶	$\mathcal{H}(x, y)$	目标边缘区域	$CN$	复高斯分布
$\rho$	灰度阶的概率	$QB$	量子平衡	$Sta$	目标状态
$E$	灰度能量	$o(X, Y)$	视觉量子几何中心	$FER$	预测错误率

### 2.1 基本定义

**定义 1(视觉场空间).** 由  $\alpha \times \beta$  个视觉采样单元组成的二维图像感光空间定义为视觉场空间(visual field space, 简称 VFS), 记为

$$VFS_t = \begin{bmatrix} V_{1,1} & V_{1,2} & \cdots & V_{1,\beta} \\ V_{2,1} & V_{2,2} & \cdots & V_{2,\beta} \\ \vdots & \vdots & \ddots & \vdots \\ V_{\alpha,1} & V_{\alpha,2} & \cdots & V_{\alpha,\beta} \end{bmatrix} \quad (1)$$

其中,  $VFS$  为视觉场空间;  $t$  为时间;  $V_{\alpha\beta}$  为视觉场空间的视觉采样单元,  $1 \leq \alpha < +\infty, 1 \leq \beta < +\infty$ . 视觉场空间模型如图 3 所示(目标外部黑色圆为背景采样单元, 目标边缘处红色圆为包含背景与前景信息的边缘采样单元, 目标内部黄色圆为前景采样单元).



**Fig.3** Visual field space model

图 3 视觉场空间模型

本文提出的“视觉场空间”与计算摄像学领域中的“视觉场”不同,本文“视觉场空间”是利用多个视觉采样单

元对运动目标空间信息进行分类和建模,主要包括目标外部的背景视觉单元、目标边缘处的边缘视觉单元以及目标内部的前景视觉单元这 3 个类别,而计算摄影学领域的“视觉场”概念不具有这种分类和建模特点.

**定义 2(视觉量子).** 视觉场空间中包含目标背景信息、前景信息和背景与前景交界处信息的视觉采样单元定义为视觉量子(vision quantum,简称 VQ),记为

$$VQ_{\alpha,\beta}(X,Y,D,P,T) \tag{2}$$

其中, $X,Y$ 为视觉量子的几何中心坐标, $D$ 为视觉量子的直径, $P$ 为视觉量子的量子频率(详见定义 4), $T$ 为时间,下标 $\alpha,\beta$ 表示视觉量子在视觉场空间中的矩阵位置是 $\alpha$ 行 $\beta$ 列.

本文提出的“视觉量子”与量子力学中“量子”不同,本文“视觉量子”是模仿生物视觉系统的“视觉单元”进行计算机视觉研究的图像信息的最小载体,而量子力学中的“量子”是物理学领域中研究微观粒子能量变化的最小单位和能量载体.

根据视觉量子包含信息不同,将视觉量子分为 3 类,分别为包含目标背景信息的背景视觉量子、包含目标前景信息的前景视觉量子、包含目标背景与前景边界信息的边缘视觉量子,分别如图 4~图 6 所示.

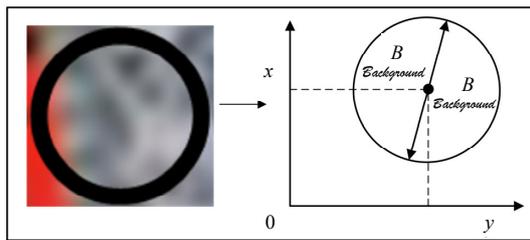


Fig.4 Background visual quantum

图 4 背景视觉量子

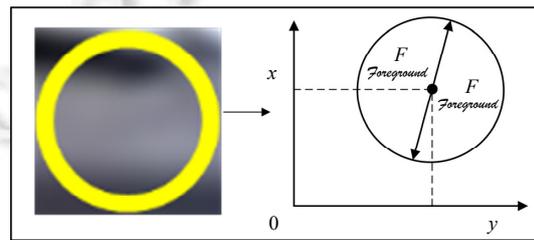


Fig.5 Foreground visual quantum

图 5 前景视觉量子

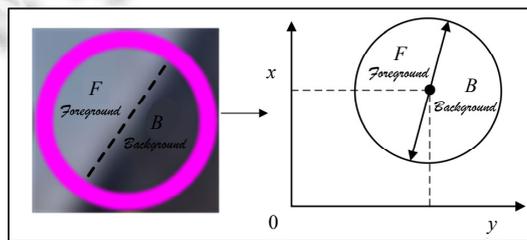


Fig.6 Edge visual quantum

图 6 边缘视觉量子

由图 4 和图 5 可知,背景视觉量子内部的背景信息和前景视觉量子内部的前景信息分布较为均匀,对于变结构运动目标不具有跟踪意义.由图 6 可知,边缘视觉量子既包含了目标前景信息,又包含了目标背景信息,对于变结构运动目标具有较好的跟踪意义.但是背景与前景交界处信息分布较为模糊,难以准确区分,由于信息熵可以有效地计算图像信息的分布情况,而灰度的能量可以有效地增强相似灰度信息的差别,因此,为了有效地提取和判别视觉量子内部信息的类别和分布情况,提出了熵能量的概念.

**定义 3(熵能量).** 在视觉场空间中,对 $\forall VQ_{\alpha,\beta} \in VFS$ ,存在 $VQ_{\alpha,\beta}$ 的前景信息熵 $\zeta^F$ 的概率分布和背景信息熵 $\zeta^B$ 的概率分布,将信息熵与单位灰度能量的乘积定义为 $VQ_{\alpha,\beta}$ 的熵能量(entropy energy,简称 EE),记为

$$EE_{\alpha,\beta}^{B|F} = \gamma(E_{\alpha,\beta}, S_{\alpha,\beta}) \sum_{\kappa=1}^{\sigma} \sum_{\lambda=1}^{\tau} (\zeta_{\kappa|\lambda}^{B|F})^2 \tag{3}$$

其中, $B|F$ 指背景 $B$ 或前景 $F$ , $\kappa|\lambda$ 指 $\kappa$ 或 $\lambda$ ,下标 $\alpha,\beta$ 表示视觉量子 $VQ_{\alpha,\beta}$ 在视觉场空间中的矩阵位置是 $\alpha$ 行 $\beta$ 列,信息熵 $\zeta_{\kappa|\lambda}^{B|F} = A^{B|F} \rho_{\kappa|\lambda} \log_2 \rho_{\kappa|\lambda}$ , $A^F$ 为 $VQ_{\alpha,\beta}$ 内前景信息熵分布的极大值所对应的灰度阶, $A^B$ 为 $VQ_{\alpha,\beta}$ 内背景信息熵分布的极大值所对应的灰度阶, $\rho_{\kappa}$ 为 $A^B$ 出现的概率, $\rho_{\lambda}$ 为 $A^F$ 出现的概率, $\sigma,\tau$ 为 $VQ_{\alpha,\beta}$ 内背景信息熵与前景信息

熵的数量,  $\zeta_k^B \in [\zeta_1^B, \zeta_2^B, \dots, \zeta_\sigma^B], \zeta_\lambda^F \in [\zeta_1^F, \zeta_2^F, \dots, \zeta_\tau^F], E_{\alpha,\beta}$  为灰度能量信息,  $E_{\alpha,\beta} = \sum G(x,y)^2$ ,  $G(x,y)$  为视觉量子内部像素点  $(x,y)$  的灰度值,  $S_{\alpha,\beta}$  为  $VQ_{\alpha,\beta}$  的几何面积,  $S_{\alpha,\beta} = \pi D^2/4$ , 则  $VQ_{\alpha,\beta}$  的单位灰度能量为

$$\gamma(E_{\alpha,\beta}, S_{\alpha,\beta}) = \frac{E_{\alpha,\beta}}{S_{\alpha,\beta}} = \frac{\sum G(x,y)^2}{\pi D^2/4}$$

熵能量函数波形图如图 7 所示, 图中纵坐标为熵能量, 横坐标为熵的数量, 函数振幅表明视觉量子中边缘信息变化的剧烈程度, 其中, 1(25,9) 表示图 3 中第 1 幅图像的第 25 行第 9 列视觉量子的熵能量波形图, 2(21,11) 表示图 3 中第 2 幅图像的第 21 行第 11 列视觉量子的熵能量波形图, 4(16,6) 表示图 3 中第 4 幅图像的第 16 行第 6 列视觉量子的熵能量波形图, 5(21,7) 表示图 3 中第 5 幅图像的第 21 行第 7 列视觉量子的熵能量波形图。

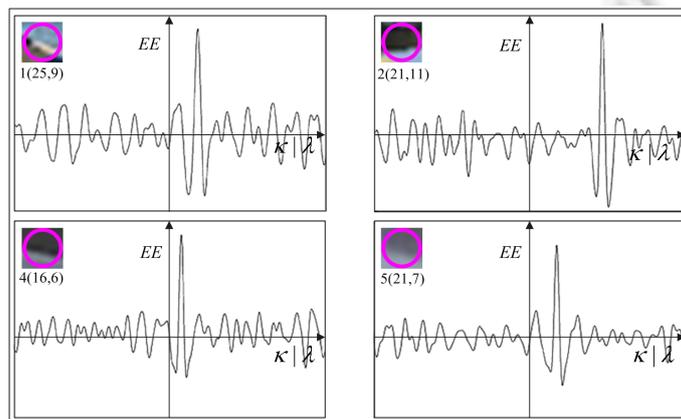


Fig.7 Entropy energy waveform

图 7 熵能量波形图

熵能量可以有效地提取和判别视觉量子内部信息的类别及其分布情况, 但由于信息熵  $\zeta^{BF}$  的概率分布集合中存在多个概率值, 熵能量  $EE^{BF}$  的函数波形走向呈不规则曲线状态. 因此, 对熵能量  $EE^{BF}$  进行二维核变换, 然后进行滤波处理, 提取视觉量子  $VQ$  内背景和前景的频率信息.

定义 4(量子频率). 对熵能量  $EE^{BF}$  的二维核变换定义为  $VQ$  的量子频率(quantum frequency, 简称 QF), 记为

$$P^{BF}(u, v, t) = \int_0^{\kappa} \int_0^{\lambda} \int_0^t EE^{BF} e^{j2\pi(u\lambda/\kappa + v\gamma/\lambda)t} dx dy dt \quad (4)$$

化简得:

$$P^{BF}(u, v, t) = EE^{BF} \kappa \lambda / 4\pi^2 uvt [e^{j2\pi t(u\lambda/\kappa + v\gamma/\lambda)} + e^{j2\pi u\lambda/\kappa} + e^{j2\pi v\gamma/\lambda} + 1] \quad (5)$$

其中,  $P^{BF}$  是指背景量子频率  $P^B$  或前景量子频率  $P^F$ ,  $e^{j2\pi t(u\lambda/\kappa + v\gamma/\lambda)}$  为二维变换核,  $e^{j2\pi(u\lambda/\kappa + v\gamma/\lambda)} + e^{j2\pi u\lambda/\kappa} + e^{j2\pi v\gamma/\lambda} + 1$  称为量子频率变换核的组合,  $EE^{BF} \kappa \lambda / 4\pi^2 uvt$  为变换系数.

由图 7 可以看出, 熵能量是围绕在目标边缘处的离散波动效应. 若将量子频率的变换系数进行归一化处理, 可以有效滤除噪音和杂波干扰, 提高目标量子频率提取的准确性. 归一化操作定义如下:

$$\int_{-\infty}^{+\infty} |EE^{BF} \kappa \lambda / 4\pi^2 uvt| dt = 1 \quad (6)$$

在目标运动过程中, 对目标前景和背景交界处的边缘视觉量子信息的提取和分析可以有效地跟踪目标位置, 因此, 对目标边缘信息的提取变得尤为重要. 由数学理论可知, 二元函数的二阶偏导的物理意义是加速度变化快慢的量, 其中, 加速度是瞬时速度的变化量, 而量子频率瞬时变化的快慢表达了图像边缘梯度信息变化的剧烈程度. 因此, 可以通过计算局部量子频率的二阶偏导来确定目标边缘区域  $\mathcal{R}(x,y)$ , 则

$$\mathcal{R}(x, y) = \frac{\partial^2 P}{\partial u^2} + \frac{\partial^2 P}{\partial v^2} + \frac{\partial^2 P}{\partial u \partial v} + \frac{\partial^2 P}{\partial v \partial u} \quad (7)$$

由目标运动可微分定理,函数  $P$  在任意点  $(u,v)$  是可微分的,则有:

$$\frac{\partial P}{\partial \vec{T}} = \frac{\partial P}{\partial u} \cos \theta + \frac{\partial P}{\partial v} \sin \theta \quad (8)$$

其中,  $\vec{T}$  为相邻极值间的向量.因此,目标边缘区域  $\mathcal{R}(x,y)$  为

$$\mathcal{R}(x,y) = \frac{\partial^2 P}{\partial u^2} \cos^2 \theta + \frac{\partial^2 P}{\partial v^2} \sin^2 \theta + \left( \frac{\partial^2 P}{\partial u \partial v} + \frac{\partial^2 P}{\partial v \partial u} \right) \cos \theta \sin \theta \quad (9)$$

令  $\theta$  取值  $0, \pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6$ .  $\vec{T}$  和  $\theta$  如图 8 所示.

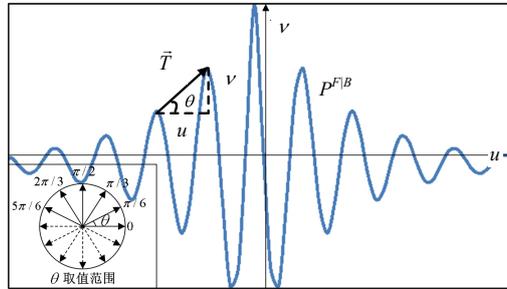


Fig.8 Sketch map of  $\vec{T}$  and  $\theta$   
图 8  $\vec{T}$  和  $\theta$  示意图

对于快速变结构运动目标而言,其边缘视觉量子内包含的边界信息均为不规则的非线性信息,给统计相邻时刻视觉量子内部边缘信息的变化幅度和变化方向都带来较大困难.因此,将边缘视觉量子内部非线性边界近似为线性边界进行统计和计算,可以消除不规则非线性边缘信息对视觉量子分析的影响,近似过程如图 9 所示.

在近似状态下,假设目标由  $t$  时刻运动到  $t+1$  时刻,则视觉量子采集的目标边缘变化过程如图 10 所示,对应的量子频率变化过程如图 11 所示.

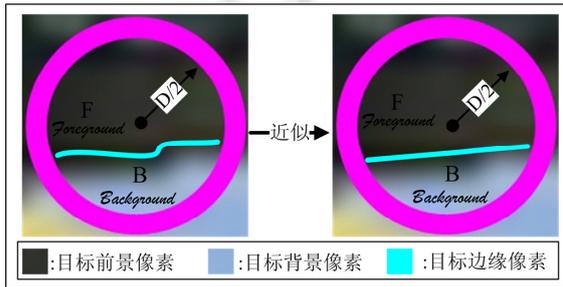


Fig.9 Approximate representation of irregular object edge  
图 9 不规则目标边缘的近似表示

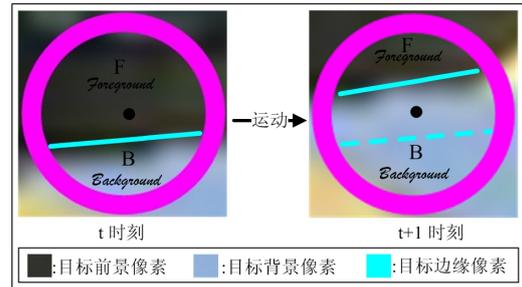


Fig.10 Change of object edge  
图 10 目标边缘变化过程

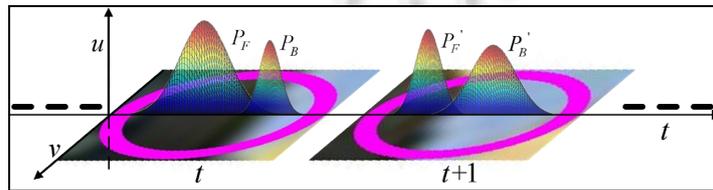


Fig.11 Sketch map of quantum frequency and its changing process between time  $t$  to  $t+1$   
图 11  $t \sim t+1$  时刻量子频率及其变化过程示意图

由图 10,视觉量子中的目标边缘由  $t$  时刻变化到  $t+1$  时刻,即由虚线位置跳变至实线位置,所对应的量子频

率跳变过程如图 11 所示,即  $P^F \rightarrow P^{F'}$  且  $P^B \rightarrow P^{B'}$ . 其中,  $P^F, P^{F'}$  与  $P^B, P^{B'}$  是图像采集设备采集的相邻两帧图像的量子频率计算值,属于测量的状态值. 而实际上,根据目标运动可微分定理,  $P^F \rightarrow P^{F'}$  的峰值之间是连续递变的;同理,  $P^B \rightarrow P^{B'}$  的峰值之间也是连续递变的. 因此,为了有效刻画量子频率峰值实际的递变过程,需要对目标前景与背景的量子频率峰值的测量状态进行理想化. 即将量子频率的跳变过程进行连续化描述. 量子频率峰值连续递变过程理想化状态如图 12 所示,其中,左侧图的实线为量子频率的测量状态,右侧图的实线为量子频率理想化之后的实际状态.

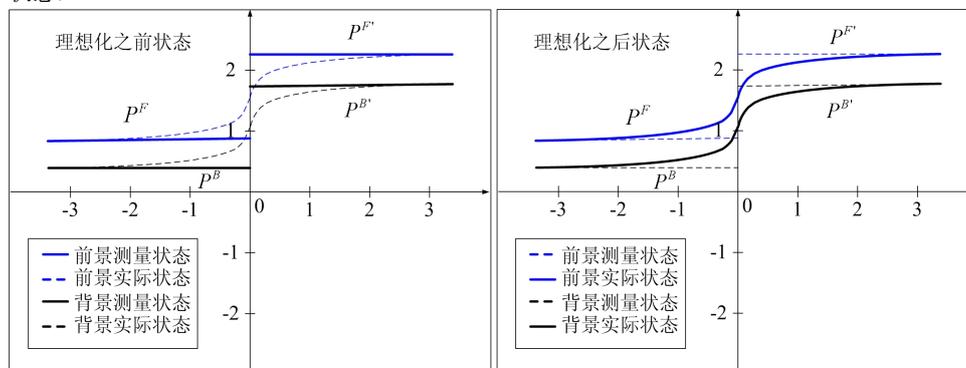


Fig. 12 Idealized statuses in the continuous variation process of quantum frequency's peaks

图 12 量子频率峰值连续递变过程理想化状态

**性质 1(阶跃不变性).** 连续的视频采集实际上是一种间歇性的图像抽样且抽样过程是均匀的,根据目标运动可微分定理及图 12 可知,当连续帧间的时间间隔趋于 0 时,任一相邻时刻状态下的目标边缘视觉量子的量子频率对微小的时间  $t$  的积分相等. 将这种性质定义为量子频率的阶跃不变性,记为

$$\lim_{\Delta t \rightarrow 0} \frac{\int_{t-\Delta t}^t P \in P^F \cap P \in P^B}{\int_{t+1-\Delta t}^{t+1} P' \in P^{F'} \cap P' \in P^{B'}} \equiv 1 \quad (10)$$

其中,  $P$  为  $t$  时刻量子频率,  $P'$  为  $t+1$  时刻量子频率,  $\Delta t$  为连续相邻两帧间的时间间隔.

证明: 当  $\Delta t \rightarrow 0$  时, 视频采集帧率  $FPS \rightarrow +\infty$ . 此时, 相邻时刻  $t$  与  $t+1$  的量子频率  $P^F \rightarrow P^{F'}$  且  $P^B \rightarrow P^{B'}$ .  $\square$

由图 11 可知,  $P \in P^F \cap P \in P^B$  为目标前景与背景交界处的量子频率范围, 在目标运动过程中具有阶跃不变性, 对于视觉量子中的边缘信息的跳变速度具有很好的鲁棒性, 同样, 对于前景和背景区域的光照变化、颜色变化以及目标遮挡有很好的抗干扰特点.

**定义 5(量子平衡).** 视觉量子  $VQ_{\alpha, \beta}(X, Y, D, P, T)$  沿向量  $\vec{\delta}$  移动, 使视觉量子几何中心  $o(X, Y)$  与目标边缘区域几何中心  $o'(X', Y')$  重合, 将此时的状态定义为量子平衡(quantum balance, 简称 QB), 记为

$$QB \{P^F = P^B, o(X, Y) = o'(X', Y')\} \quad (11)$$

其中,  $P^F$  为目标前景量子频率,  $P^B$  为目标背景量子频率,  $o(X, Y)$  为视觉量子几何中心坐标,  $o'(X', Y')$  为目标边缘区域几何中心坐标. 设目标边缘区域  $\mathcal{M}(x, y)$  的像素点数量为  $n$ ,  $x_i, y_i$  为  $\mathcal{M}(x, y)$  的像素点坐标,  $1 \leq i \leq n$ , 则有:

$$X' = \sum_{i=1}^n x_i / n, Y' = \sum_{i=1}^n y_i / n.$$

移动向量  $\vec{\delta}$  是以  $o(X, Y)$  为起点、 $o'(X', Y')$  为终点的向量, 即:

$$\vec{\delta} = \overrightarrow{oo'} = (X' - X, Y' - Y) \quad (12)$$

本文提出的量子平衡与量子医学领域中的量子平衡不同, 本文提出的量子平衡是指视觉量子沿向量移动后与目标边缘几何中心重合的状态, 而量子医学领域中的量子平衡是指人体细胞内部电子运动由异常恢复正常的状态.

图 10 中,  $t$  和  $t+1$  时刻的视觉量子达到量子平衡状态如图 13 所示, 图 13 中第 2 列  $t+1$  时刻视觉量子移动至

量子平衡状态时的量子频率变化如图 14 所示.

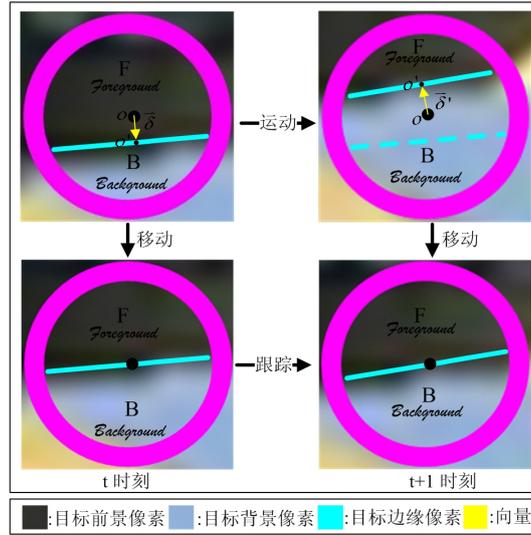


Fig.13 Status of quantum balance

图 13 视觉量子平衡状态

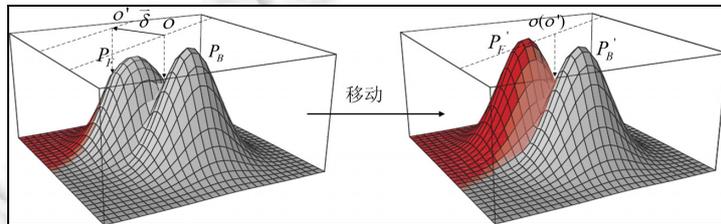


Fig.14 Quantum frequency variation in the process of an object moving to quantum balance status

图 14 量子平衡移动过程中量子频率变化过程

由图 13 可见,视觉量子经过移动达到量子平衡状态时,可以对快速变结构目标运动的局部边缘信息进行有效跟踪.若在视频序列初始化中选待跟踪的目标前景信息,则可以通过多个视觉量子对目标准确定位和跟踪.

定义 6(量子簇). 将达到量子平衡状态的  $N$  个边缘视觉量子的组合定义为量子簇(quantum cluster,简称 QC),记为

$$QC^{QB}(t) = [VQ_1^{QB}(t), VQ_2^{QB}(t), \dots, VQ_N^{QB}(t)]^T,$$

其中,  $N$  为量子簇中边缘视觉量子的数量,  $N \geq 3$ ;  $QB$  为量子平衡状态;  $VQ_k^{QB}(t)$  为  $t$  时刻达到量子平衡状态的视觉量子,  $3 \leq k \leq N$ . 对于任意  $VQ_k^{QB}(t)$ , 存在  $VQ_k^{QB}(t) \in QC^{QB}(t) \in VFS_t$ . 量子簇如图 15 所示.



Fig.15 Sketch map of a quantum cluster

图 15 量子簇示意图

## 2.2 目标状态预测

假设在视频图像内向目标区域辐射  $M$  个视觉量子  $(x_k, y_k), k=1, \dots, M$ , 经移动达到量子平衡后, 得到  $N$  个候选的视觉量子  $(x_l, y_l), l=1, \dots, N$ , 则第  $k$  个视觉量子被候选为第  $l$  个视觉量子的概率分布可被看作熵能量与复高斯随机变量  $\alpha_{kl}$  的乘积, 且有:

$$\alpha_{kl} \sim CN(0, \varepsilon_{kl}^2) \quad (13)$$

其中, 复高斯分布可参考文献[38].  $\varepsilon_{kl}^2$  为复高斯随机变量与均值偏离的程度,  $\alpha_{kl}$  的变化反映了扩大对目标区域辐射视觉量子导致候选视觉量子增加的累积效果.

候选的视觉量子中包含了一部分无效的视觉量子, 它们是由相似运动目标干扰而形成量子平衡的视觉量子. 因相似目标的出现和运动对整个视觉场空间具有随机性和均匀性, 其概率密度满足正态的均匀分布. 将其对目标的作用看作一组复高斯白噪声干扰过程, 则作为候选的视觉量子被复高斯白噪声  $\beta_l(t)$  影响的概率分布为

$$\beta_l(t) \sim CN(0, \phi_0) \quad (14)$$

其中,  $\phi_0$  为受复高斯白噪声影响的候选视觉量子与均值的偏离程度.

假设候选视觉量子整体对目标跟踪是有效的, 则可设计似然函数以准确预测目标状态信息  $Sta = (x, y, \bar{v})^T$ , 其中包括目标位置  $(x, y)$  和目标运动速度  $\bar{v}, \bar{v} = \bar{v}_x + \bar{v}_y$ . 将候选视觉量子看作一个  $N \times 1$  维向量组合:

$$VQ(t) = [VQ_1(t), VQ_2(t), \dots, VQ_M(t)]^T \quad (15)$$

则目标状态  $Sta$  的极大似然估计为

$$Sta = \arg \max_{Sta} \sum_{k=1}^M \sum_{l=1}^N [\varphi_{kl} / \phi_0 (\varphi_{kl} + 1)] \times |VQ_{kl}(EE^{BIF}(Sta))|^2 \quad (16)$$

其中,

$$EE^{BIF}(Sta) = \gamma(E, S) \int_{-\infty}^{\infty} VQ(t) A^{BIF} \rho_{\kappa|\lambda}(Sta) \times \log_2 \rho_{\kappa|\lambda}(Sta) dt \quad (17)$$

$\varphi_{kl}$  为  $VQ(t)$  的信噪比:

$$\varphi_{kl} = \frac{\rho_{\kappa|\lambda} \varepsilon_{kl}^2}{\phi_0} \quad (18)$$

若给定一组标准正交基  $\{\zeta_1(t), \zeta_2(t), \dots, \zeta_k(t)\}$ , 则  $VQ_l(t)$  可拓展为如下形式:

$$VQ_l(t) = \lim_{K \rightarrow \infty} \sum_{k=1}^K \partial_{kl} \xi_k(t) \quad (19)$$

其中,  $\partial_{kl}$  是第  $k$  个正交基函数的系数, 且  $\partial_{kl} = \int_{-\infty}^{\infty} VQ_l(t) \xi_k^*(t) dt$ ;  $(\cdot)^*$  为复共轭操作. 因此有:

$$\beta_{kl} \sim CN(0, \phi_0) \quad (20)$$

$$EE^{BIF} \sim CN(0, \rho_{\kappa|\lambda} \varepsilon_{kl}^2 + \phi_0) \quad (21)$$

由于似然函数与似然比成正比例关系, 因此当目标图像中存在前景信息  $H_1$  和噪声干扰信息  $H_0$  且  $k > M$  时, 根据  $\beta_{kl}$  与  $EE^{BIF}$  的概率分布,  $VQ_l$  的似然比可以派生为

$$\Gamma(VQ_l | Sta, H_1) \propto \frac{\Gamma(VQ_l | Sta, H_1)}{\Gamma(VQ_l | H_0)} \quad (22)$$

其中,

$$\begin{aligned} \Gamma(VQ_l | Sta, H_1) &\propto \frac{\Gamma(VQ_l | Sta, H_1)}{\Gamma(VQ_l | H_0)} = \prod_{k=1}^M \frac{\Gamma(VQ_{kl} | Sta, H_1)}{\Gamma(VQ_{kl} | H_0)} \prod_{k=M+1}^{\infty} \frac{\Gamma(VQ_{kl} | H_1)}{\Gamma(VQ_{kl} | H_0)} \\ &= \prod_{k=1}^M \frac{\phi_0}{\rho_{\kappa|\lambda} \varepsilon_{kl}^2 + \phi_0} \exp \left\{ \frac{\rho_{\kappa|\lambda} \varepsilon_{kl}^2 |VQ_{kl}|^2}{\phi_0 (\rho_{\kappa|\lambda} \varepsilon_{kl}^2 + \phi_0)} \right\} \prod_{k=M+1}^{\infty} \frac{\Gamma(\beta_{kl})}{\Gamma(\beta_{kl})} \\ &= \prod_{k=1}^M \frac{1}{1 + \varphi_{kl}} \exp \left\{ \frac{\varphi_{kl} |\gamma(E, S) \int_{-\infty}^{\infty} VQ(t) A^{BIF} \rho_{\kappa|\lambda}(Sta) \log_2 \rho_{\kappa|\lambda}(Sta) dt|^2}{\phi_0 (\varphi_{kl} + 1)} \right\} \end{aligned}$$

因为候选视觉量子之间是非相关的,则  $VQ(t)$  的似然函数可以写成如下形式:

$$\Gamma(VQ(t)|Sta, H_1) \propto \prod_{l=1}^N \prod_{k=1}^M \frac{1}{1+\varphi_{kl}} \times \exp \left\{ \frac{\varphi_{kl} \left| \gamma(E, S) \int_{-\infty}^{\infty} VQ(t) A^{B/F} \rho_{\kappa|\lambda}(Sta) \log_2 \rho_{\kappa|\lambda}(Sta) dt \right|^2}{\phi_0(\varphi_{kl}+1)} \right\} \quad (23)$$

对  $\Gamma(VQ|Sta, H_1)$  取对数, 则其对数似然函数可以写成:

$$\ln \Gamma(VQ(t)|Sta, H_1) = \sum_{k=1}^M \sum_{l=1}^N \left\{ \frac{\varphi_{kl} \left| \gamma(E, S) \int_{-\infty}^{\infty} VQ(t) A^{B/F} \rho_{\kappa|\lambda}(Sta) \log_2 \rho_{\kappa|\lambda}(Sta) dt \right|^2}{\phi_0(\varphi_{kl}+1)} \right\} + C \quad (24)$$

其中,  $C$  是一个常数. 由于  $\Gamma(VQ(t)|Sta, H_1)$  的导数存在, 令其导数为 0, 解方程并验证方程的解, 使似然函数  $\ln \Gamma(VQ(t)|Sta, H_1)$  达到最大值. 因为计算过程对于任何观测样本值都成立, 因此可以代入样本值, 就会得到目标状态预测信息  $Sta = (x, y, \bar{v})^T$  的极大似然估计值.

### 2.3 算法步骤

视觉量子跟踪方法步骤描述如下:

1. 目标初始化: 在待跟踪目标图像中选定目标大致的前景范围和背景范围, 统计前景区域与背景区域的观测值. 如图 16 所示.



Fig.16 Object initialization (light line indicates foreground, dark line indicates background)

图 16 目标初始化(浅色线为前景区域,深色线为背景区域)

2. 量子频率计算:(1) 在目标初始化区域内自上而下辐射视觉量子  $VQ_{\alpha, \beta}$ , 形成视觉场空间  $VFS$ ; (2) 计算并统计视觉量子内前景信息熵分布的极大值所对应的灰度阶  $A^F$  以及背景信息熵分布的极大值所对应的灰度阶  $A^B$ ; (3) 以观测值统计  $A^{B/F}$  的分布概率  $\rho_{\kappa|\lambda}$ , 计算  $VQ$  的熵能量  $EE^{B/F}$  (见公式(3)); (4) 计算  $VQ$  的量子频率  $QF$ , 将量子频率系数归一化, 滤除系统噪音和杂波干扰 (见公式(5)、公式(6)).

3. 量子平衡跟踪:(1) 计算量子频率的二阶偏导, 确定目标边缘区域  $\mathcal{M}(x, y)$  (见公式(7)~公式(9)); (2) 保留目标边缘处的视觉量子, 删除目标其他区域的视觉量子; (3) 计算以视觉量子几何中心  $o(X, Y)$  为起点、目标边缘区域几何中心  $o'(X', Y')$  为终点的向量  $\bar{\delta}$ ; (4) 沿向量  $\bar{\delta}$  将视觉量子移动至量子平衡状态; (5) 计算视觉量子的频率积分 (见公式(10)~公式(12)).

4. 目标状态预测:(1) 对目标运动状态进行极大似然估计, 预测下一帧目标位置; (2) 采集下一帧图像, 利用频率阶跃不变性, 移动视觉量子至平衡状态, 若目标因完全遮挡丢失, 则对预测位置辐射视觉量子; (3) 以预测位置的前景信息和背景信息更新  $A^{B/F}$ .

5. 输出跟踪结果: 以达到量子平衡状态的视觉量子作为目标位置信息, 输出目标位置  $\left( \frac{\sum X}{\alpha + \beta}, \frac{\sum Y}{\alpha + \beta} \right)$ .

6. 循环跟踪: 以第 5 步中输出的跟踪位置作为下一帧图像的目标初始化区域, 转至步骤 2 循环执行.

### 2.4 实例测试

为了验证视觉量子目标跟踪方法的实用性, 在复杂条件下的实际现场中对运动目标进行跟踪测试, 跟踪对象选用不同光照、不同拍摄角度、不同旋转尺度、不同形状变化的变结构机动目标. 通过视觉量子局部采样可以看出, 利用频率阶跃不变性检测变结构目标信息基本覆盖了目标关键特征区域, 抑制了非关键信息, 通过视觉

量子的频率阶跃不变性质实现了从变化的目标信息中提取不变信息的仿生视觉跟踪方法。

在复杂条件下,当运动目标出现快速的变结构运动时,目标本身的形状及尺度变化速度较快,特征变化剧烈.尤其是在动态背景下,目标前景与背景运动相似,难以进行区分,而且目标混合边缘区域的增长也为目标跟踪带来较大干扰.但基于视觉量子的目标跟踪方法在采用频率阶跃不变性检测、量子平衡跟踪和目标状态预测情况下,对目标混合边缘信息的跟踪保持了较强的稳定性和很好的鲁棒性能;同时,由于视觉量子方法对目标实行跟踪,不需要对目标多种状态下的模板信息进行计算、匹配及存储,也不需要视频图像整体进行目标运动信息计算和分割,因此,视觉量子目标跟踪方法处理的数据量较低,使该方法具有较高的跟踪速度和较强的适应能力.与现阶段著名的 TLD 方法、OLSL 方法及 WLMC 方法相比,视觉量子跟踪方法在实时性、稳定性和准确性方面具有较高的优势.

在测试中,选取了一组实际应用场景中对现有目标跟踪方法极具挑战性的高速变结构战机飞行视频数据(视频数据量为 30 884 帧,宽高为 646×468 像素,帧容量为 872KB,位深度为 24 位).对基于视觉量子理论的目标跟踪方法进行测试的结果及其跟踪过程中移动向量围绕量子中心波动的效果分别如图 17 和图 18 所示.



Fig.17 Results of maneuvering object tracking under 8 variable structure moving statuses

图 17 8 种变结构运动状态下的机动目标跟踪结果

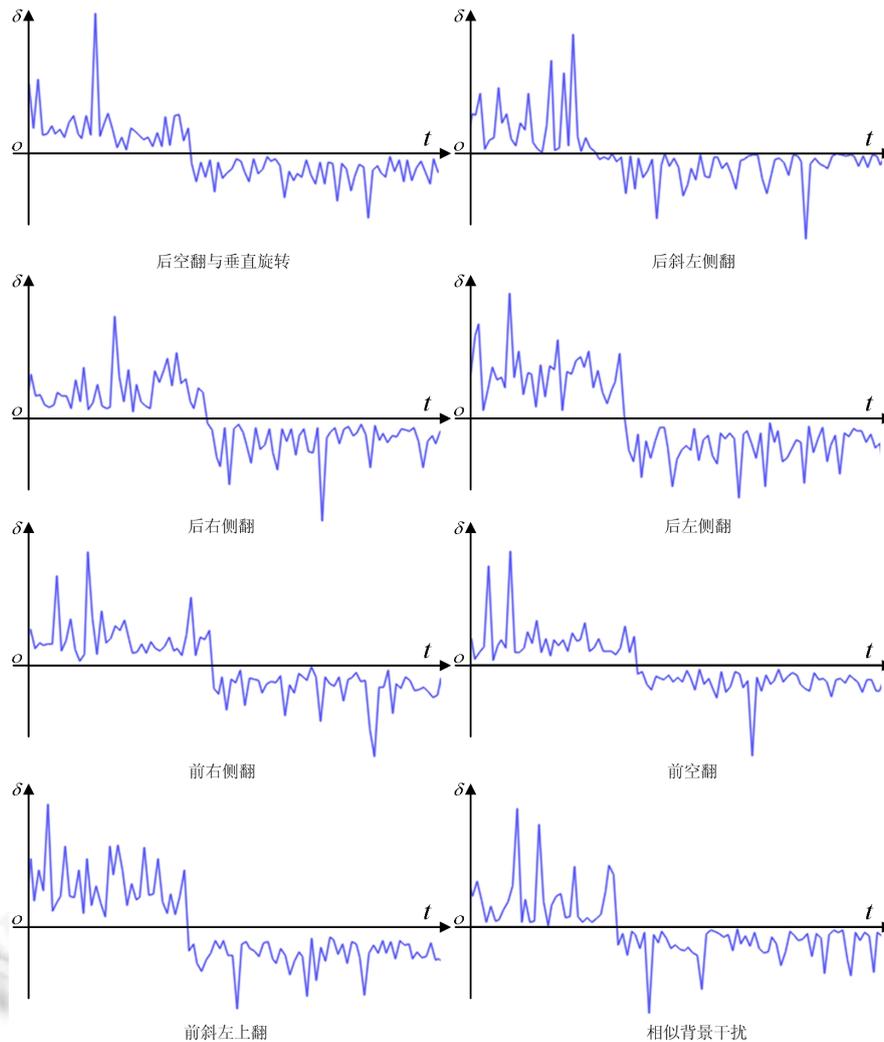


Fig.18 Fluctuating effect of moving vector around the quantum center

图 18 移动向量围绕量子中心波动效果

图 17 中浅色框图为跟踪结果,视频处理速度为 65 帧/秒,准确率为 90.7%.实例测试的目标运动状态包括:后空翻与垂直旋转、后斜左侧翻、后右侧翻、后左侧翻、前右侧翻、前空翻、前斜左上翻及相似背景干扰.

### 3 实验与相关工作对比分析

本文以视觉量子理论为核心的变结构目标跟踪方法,在 Pentium 3.06GHz/1GB 机器上采用 C++语言编程实现.为了突出视觉量子跟踪方法的优点及与相关跟踪方法的不同之处,本文介绍了目前较为流行的目标跟踪方法的实现步骤.在评估中,首先采用与本文算法应用条件相同的 IBM 多目标遮挡数据库<sup>[39]</sup>进行算法测试,同时采用跟踪与监视性能评估工作组(performance evaluation of tracking and surveillance,简称 PETS)<sup>[40]</sup>的性能评估指标对本文算法进行性能评估.PETS 主要提供视觉监控与目标跟踪算法的评估服务,早期的 PETS 主要针对人员与车辆的运动目标检测与跟踪,现在主要进行以像素和目标为单位的统计评估并且提供在线监控与跟踪算法的性能评估服务.同时也采用了美国 California 大学 Yi 等人提出的视觉跟踪评估体系<sup>[29]</sup>,对本文方法进行了公正的评价.对相关文献中作者公开发表的 Matlab 跟踪算法采用 C++语言进行了重新实现,以与本文方法进行

更为公平的对比.

### 3.1 相关工作

#### 3.1.1 TLD 跟踪方法<sup>[31]</sup>

2012年7月,英国萨里大学的学者 Kalal 提出了一种 TLD(tracking-learning-detection)单目标跟踪方法,即将跟踪-学习-检测三者进行结合,实现边跟踪、边学习、边检测.该方法是将传统的模板匹配跟踪算法、特征检测算法及 P-N 学习算法相结合,以解决目标运动过程中产生的形变和局部遮挡问题.通过 P-N 学习算法对跟踪模板的“显著特征”进行不断更新,对检测过程中的目标模型及相关参数也要进行更新,以此实现目标稳定跟踪.TLD 实现过程如图 19 所示.

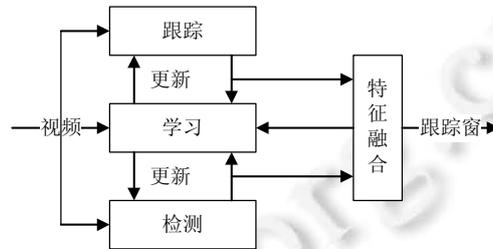


Fig.19 Implementation process of TLD method

图 19 TLD 方法实现过程

TLD 方法的缺点是,由于 TLD 方法需要不断进行模板更新和在线学习,TLD 方法只能对目标外形和尺度变化较小的目标进行跟踪,对于外形和尺度变化速度较快且出现相似背景干扰的情况,TLD 方法很容易出现失跟和跟踪错误的问题.

#### 3.1.2 OLSL 跟踪方法<sup>[33]</sup>

2013年,Yao 等学者提出了基于在线局部结构学习的跟踪算法:采用观测样本和样本数量对图像局部模板进行在线训练和分类,结合历史模型数据与样本分类对多个局部模型进行 SVM 结构化学习,建立外观模型,输出跟踪结果.当目标发生遮挡或者局部外观发生显著变化时,该帧数据不能用于增量学习,以提高模型的鲁棒性.这就要求:在每一帧,都要根据历史数据重新训练 SVM.即,为了计算当前帧的外观模型,需要将当前帧加入到在线训练与分类器中进行训练,重新计算 SVM.OLSL 算法步骤如图 20 所示.

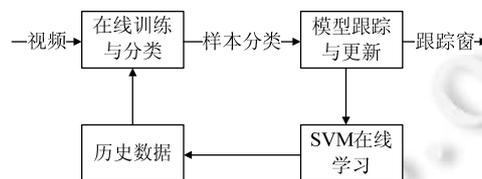


Fig.20 Implementation process of OLSL algorithm

图 20 OLSL 算法实现步骤

OLSL 方法的缺点是,OLSL 方法需要对外观模型进行反复学习和更新,同时需要对新生长的局部外观模型重新训练与分类.虽然在一定条件下显示了较好的跟踪效果,但是对于相似目标干扰和变结构目标运动,跟踪的失跟率较高,甚至容易出现跟踪错误目标的情况.

#### 3.1.3 WLMC 跟踪方法<sup>[32]</sup>

2013年,韩国首尔大学的学者 Kwon 针对突变运动问题提出了一种基于 Wang-Landau 的蒙特卡罗采样跟踪方法,以解决目标在复杂环境下的突变运动跟踪问题.该算法首先引入了 Wang-Landau 采样方法,并将其集成到一个以蒙特卡罗(MCMC)跟踪框架为基础的马尔可夫链中,通过采集当前观测样本,利用 Wang-Landau 采样

方法引入 MCMC 接受率,可有效缓解目标运动的平滑约束和实现鲁棒跟踪.WLMC 方法可对目标状态进行有效采样,对目标位置和尺度发生较大突变运动时可以进行有效的跟踪.WLMC 实现步骤如图 21 所示.

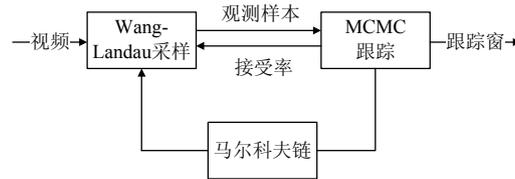


Fig.21 Implementation process of WLMC algorithm

图 21 WLMC 方法实现步骤

WLMC 方法的缺点是,WLMC 跟踪方法在蒙特卡罗(MCMC)跟踪框架的基础上,利用 Wang-Landau 采样方法引入 MCMC 接受率,可以对突变目标进行有效跟踪.但是对于长时间的连续变结构目标运动情况则不具有稳定的跟踪效果.WLMC 方法对快速的连续变结构运动目标的跟踪窗抖动性较强,稳定性较差.

### 3.2 IBM 评估指标及评估结果

目标检测的研究对象主要包括像素和目标两种,因此,目前主要有基于像素和基于目标的 2 种评估方案.基于像素级别的评估包含 4 项统计指标,分别是目标矩形框内的前景帧数量 TP(true positive)、目标矩形框外的背景帧数量 TN(true negative)、目标矩形框外的前景帧数量 FP(false positive)、目标矩形框内的背景帧数量 FN(false negative).基于以上 4 种数据指标,PETS 提出了相应的评估指标对目标跟踪算法进行性能评估,具体指标见表 2.

Table 2 Evaluation indexes of monitoring and tracking algorithms

表 2 监控与跟踪算法性能评估指标

项目名称	计算过程
TDR (tracker detect rate)	$TDR=TP/(TP+FN)$
PP (positive prediction)	$PP=TP/(TP+FP)$
FNR (false negative rate)	$FNR=FN/(TP+FN)$
FAR (false alarm rate)	$FAR=FP/(TP+FP)$
ACC (accuracy)	$ACC=(TN+TP)/NTF$
FPR (false positive rate)	$FPR=FP/(TN+FP)$

表 2 中,NTF(number of total frame)为经过算法处理的帧总数.基于像素的评估方法需要用人工对视频数据的真实场景进行准确标定,人工实现较为困难.而基于目标为单位的统计评估方法,表 2 所列出的参数也同样适用.

以目标数量进行统计的评估方法,主要通过系统的输出目标区域和矩形框标定的目标区域的重叠区域进行判别.主要判断矩形框标定目标是否处于待检测目标的边界以内:

- 若处于边界内,则判定为目标检测成功,此时, $TP=TP+1$ ;
- 否则判定为目标没有检测成功,此时, $FN=FN+1$ .

若待检测目标搜索不到可以成功匹配的矩形框,则标定此待检测目标为 FP,此时, $FP=FP+1$ ;若待检测目标与矩形框目标产生交错,则列为候选目标的像素区域,此时, $TN=TN+1$ .基于目标的评估方法简便、快捷,可以对跟踪算法的监控与跟踪效果、虚警概率进行有效的性能评估.同时,为了定量评价本文算法处理遮挡下目标状态预测能力,在此引入一个新的评价指标——预测错误率(forecast error rate,简称 FER),用于描述目标在被遮挡前后错误预测的概率(包括发生实际运行轨迹和预测轨迹出现较大偏离的情况),定义如下:

$$FER = \sum(\sim \mu_j) / \sum(\sim \psi_l),$$

其中, $\mu_j$  为第 j 次遮挡前后预测状态值, $\psi_l$  为第 l 次遮挡前后分裂状态值.当 $\mu=0$ 时,预测错误;当 $\mu=1$ 时,预测有效.当 $\psi=0$ 时,遮挡发生分裂;当 $\psi=1$ 时,遮挡未分裂.

本文的 VQ 跟踪方法、TLD 方法、OLSL 方法及 WLMC 方法在 IBM 多目标遮挡数据库的 15 组视频,共 106 032 帧图像的测试结果如图 22 所示,其中,每组实验均反复进行 20 次.综合测试结果表明,VQ 跟踪方法具有较高的跟踪性能:目标外观与背景相似干扰条件下的漏检概率 FNR 为 0.011,如图 22 所示.由于跟踪过程中充分考虑了运动目标邻域量子频率间存在的频率阶跃不变性,然后将该性质纳入量子平衡移动中,并对目标进行预测跟踪,降低了系统和环境噪声引起的虚警概率 FPR,FPR 的平均测试结果为 0.037.同时,有效解决了变结构运动及遮挡引起的目标丢失问题,跟踪准确率 TDR 平均为 0.912,预测错误率平均为 0.13.在目标跟踪过程中,由于采用了似然估计对目标状态进行预测式跟踪,使得本文算法具有较高的实时性和准确率,对 IBM 多目标遮挡数据库的处理速度平均为 75 帧/秒.

为了进一步验证本文的前趋预测跟踪方法在相似背景干扰强度、目标运动速度和多目标遮挡强度变化情况下的跟踪性能,本文在 IBM 多目标遮挡数据库的 15 组视频上进行了大量的性能测试,每组实验反复进行 20 次,以均值表示平均测试结果.同时,与 TLD 算法、OLSL 算法、WLMC 算法在相似背景干扰强度、变结构运动速度和多目标遮挡强度变化的情况下三者的跟踪性能进行比较.

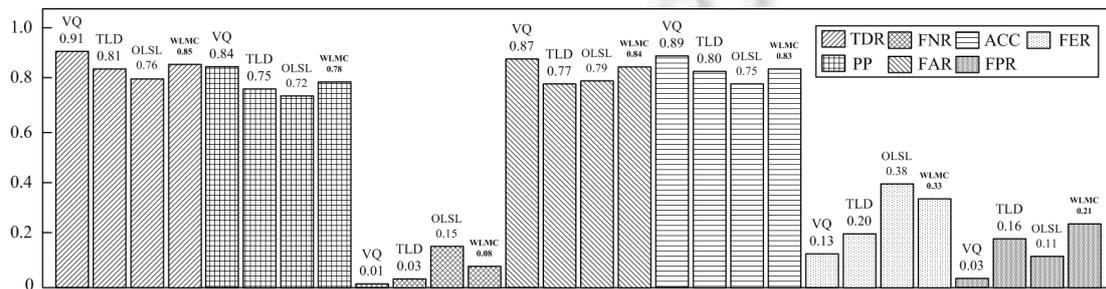


Fig.22 Assessment result of multi-objective occlusion database

图 22 多目标遮挡数据库评估结果

### 3.2.1 相似背景干扰强度变换

在相似背景干扰强度递增且检测频率不变的情况下进行跟踪误差与实时性误差实验,如图 23 所示.实验结果表明,当相似背景干扰强度逐步递增时,本文的视觉量子跟踪方法在运动目标相似背景的递增式干扰条件下的跟踪误差和实时误差都普遍较低,相对于其他 3 种方法,本文的跟踪方法对相似背景干扰条件下的目标跟踪具有显著的稳定性和实时性.

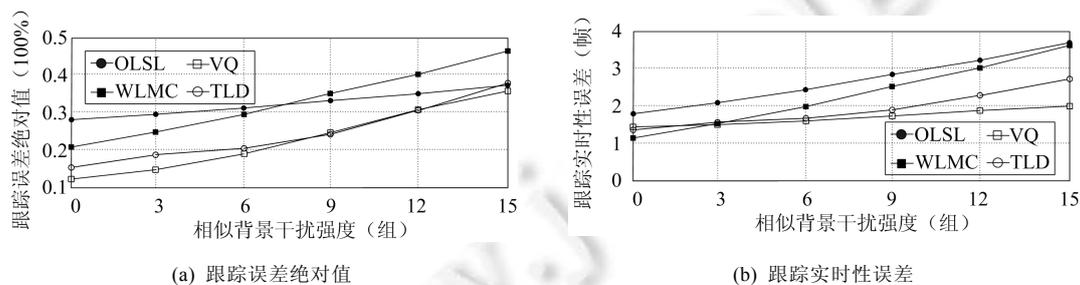


Fig.23 Performance comparison about the intensity transformation of interference from similar background

图 23 相似背景干扰强度变换的性能对比

### 3.2.2 变结构运动速度变换

在目标运动速度递增且检测频率不变的情况下进行跟踪误差与实时性误差实验,如图 24 所示.实验结果表明,当目标运动速度逐渐增大时,本文的视觉量子跟踪方法的跟踪误差递增幅度较小,实时性误差基本保持在较

低水平,验证了本文基于视觉量子的目标跟踪方法对于速度变化的运动目标跟踪具有明显优势;而其他 3 种方法对目标运动速度递增条件下的跟踪误差和实时性误差都在逐步递增,比本文 VQ 方法的跟踪误差和实时性误差平均高出 0.2 和 0.3,不具有明显的跟踪优势.

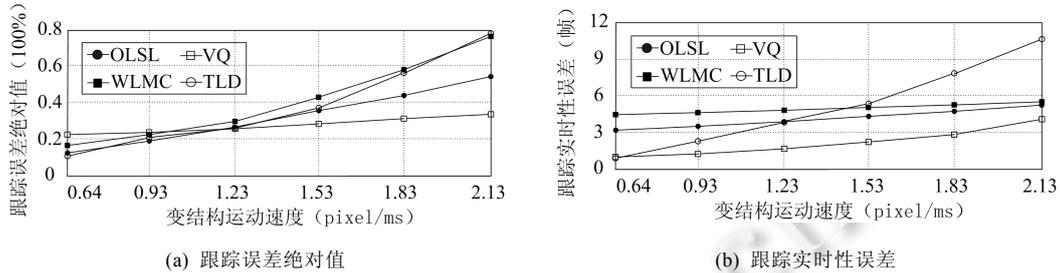


Fig.24 Performance comparison about speed transformation of object motion

图 24 目标运动速度变换的性能对比

### 3.2.3 多目标运动遮挡强度变换

多目标运动遮挡强度递增且检测频率不变的情况下进行跟踪误差与实时性误差实验如图 25 所示.从图中可以看出:当多目标交叉式运动的遮挡强度逐步增加时,本文视觉量子跟踪方法的跟踪误差保持在较低水平,增幅较小,跟踪误差较低,对多目标遮挡具有很好的鲁棒性;而其他 3 种方法跟踪误差较高,尤其是 TLD 方法,虽然在目标遮挡强度较低时保持了很好的跟踪准确性,跟踪误差最小,但是当目标遮挡强度递增时,跟踪误差增幅较大,对目标遮挡跟踪的鲁棒性较低.在实时性误差当中,4 种方法在目标遮挡强度递增时,都保持了较好的鲁棒性,但是本文 VQ 方法的跟踪实时性误差最低,比其他 3 种方法具有较高的抗遮挡跟踪优势.

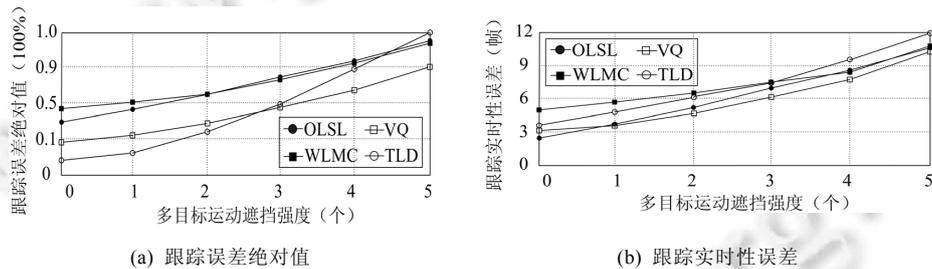


Fig.25 Performance comparison about intensity transformation of occlusion for multi-objective

图 25 多目标遮挡强度变换的性能对比

### 3.3 Benchmark 评估指标及评估结果

在 PETS 性能评估基础上,本文采用美国 California 大学 Wu 等人提出的 Benchmark 视觉跟踪评估体系,对本文方法进行了进一步公正的评价.Benchmark 评估体系提供了 50 组不同干扰条件下的视频数据库进行跟踪测试,在此基础上,提供了可以与相关算法进行对比的近 660 000 个目标跟踪窗,并注明了最常使用的跟踪序列.图 26 显示了 Benchmark 视频数据库中每个视频序列的第 1 帧进行初始化的目标窗口.视频数据库及跟踪结果可以从 <http://visual tracking.net> 下载.

Benchmark 评估方法主要通过算法可执行程序运行结果对算法性能的好坏进行定量分析与评估.评估指标包括整体跟踪性能指标 OPE、时间鲁棒性指标 TRE 和空间鲁棒性指标 SRE.

- OPE 是在第 1 帧初始化目标位置限制下的整体跟踪平均通过率.
- TRE 是指从第 1 帧跟踪到最后一帧的整体耗时与准确性.

- SRE 包含两个方面:一是程序跟踪窗中心位置与目标中心位置的平均 Euclidean 距离,另一个是程序跟踪窗边框与实际目标边框的重叠率。

最后,通过 AUC 对精度和准确率进行统计.因此,选择合理的可执行程序是进行算法性能正确评价的基础.本文根据相关文献中的算法流程并参照了作者公开发布的基于 Matlab 的跟踪程序,采用 C++ 语言对相关文献所提出的方法进行了重新实现和算法优化,以与本文方法进行更为公平的对比.

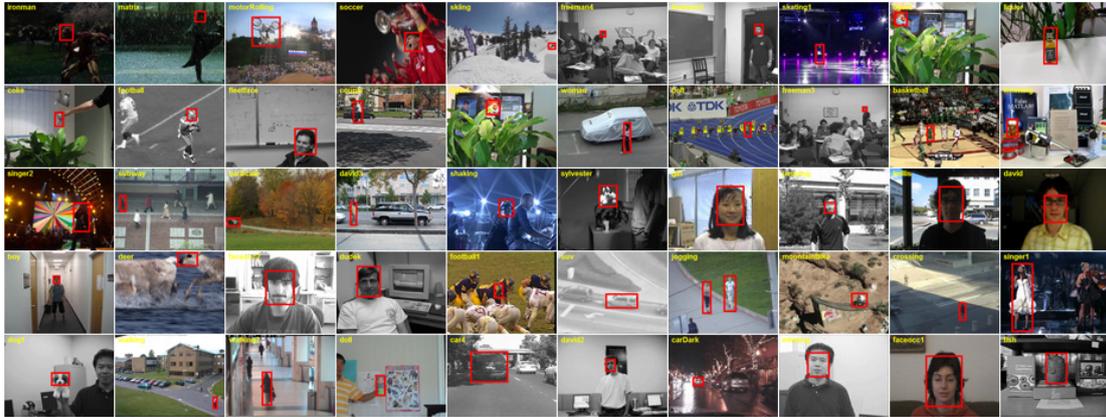


Fig.26 Samples of object initialization in Benchmark video databases

图 26 Benchmark 视频数据库目标初始化示例

在 Pentium 3.06GHz/1GB 机器上对 VQ 方法、TLD 方法、OLSL 方法、WLMC 方法进行了跟踪测试,测试结果如图 27 所示.测试中,以第 1 帧初始化的区域为跟踪目标,视频测试结果有效区间为第 2 帧至最后一帧,每组视频均反复测试 20 次,取均值作为测试结果,以提高评估准确性和公平性.综合实验结果,本文 VQ 方法在 Benchmark 评估测试中的平均跟踪准确率为 92.5%,视频处理速度平均为 73 帧/秒.

由图 27(a)~图 27(f)测试数据可以得出以下评估结果:

$$\begin{aligned} OPE^{s-r} &: TLD < OLSL < WLMC < VQ, \\ SRE^{s-r} &: TLD < OLSL < WLMC < VQ, \\ TRE^{s-r} &: TLD < OLSL < WLMC < VQ, \\ OPE^{pre} &: TLD < OLSL < WLMC < VQ, \\ SRE^{pre} &: TLD < WLMC < OLSL < VQ, \\ TRE^{pre} &: TLD < WLMC < OLSL < VQ, \end{aligned}$$

其中, $s-r$  为准确率, $pre$  为精度.由结果可以看出,

- 前 4 项评估结果一致.
- 在第 5 项和第 6 项中,WLMC 算法的 SRE 精度和 TRE 精度略小于 OLSL 算法.原因是,WLMC 算法在蒙特卡罗跟踪的样本输入过程中对输入样本实行接受率反馈,对样本系数进行动态调整,虽然在一定程度上提高了跟踪准确率,但跟踪系统的鲁棒性也略有降低.即便如此,WLMC 算法在整体性能上仍然好于 OLSL 算法.

4 种算法的综合评估结果如下:

$$TLD < OLSL < WLMC < VQ.$$

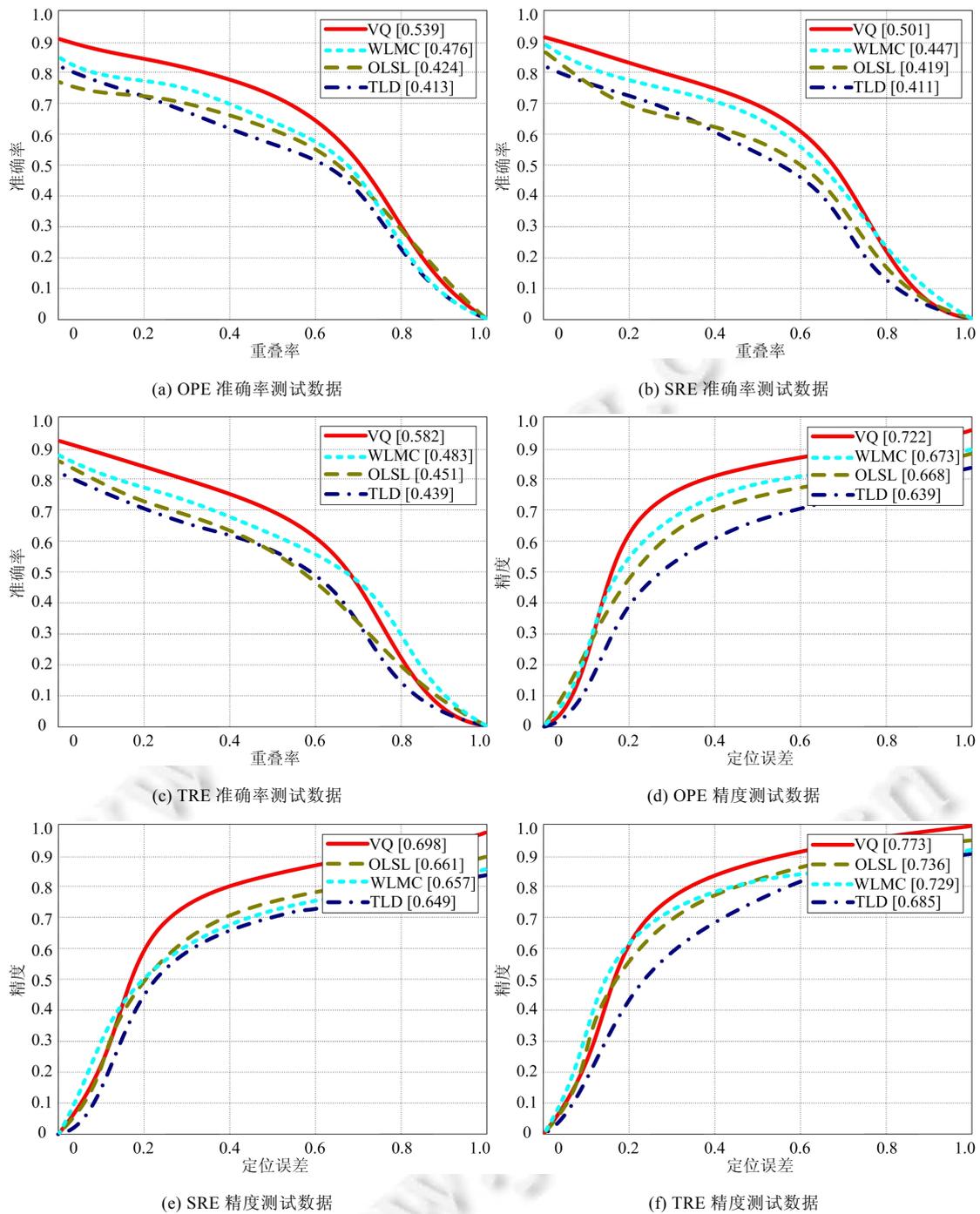


Fig.27 Benchmark performance evaluation results of VQ, TLD, OLSL and WLMC method

图 27 VQ、TLD、OLSL 和 WLMC 方法的 Benchmark 性能评估结果

在 OPE 准确率、SRE 准确率、TRE 准确率、OPE 精度、SRE 精度和 TRE 精度这 6 种评估指标的评估下,TLD 方法、OLSL 方法与 WLMC 方法显示出了较高的跟踪精度和准确率,也显示出了较好的鲁棒性能.相比之下,本

文的 VQ 跟踪方法与其他 3 种算法相比,在保持同一评测条件下,显示出了更高的跟踪精度和跟踪准确率.同时,与其他 3 种方法相比,视觉量子跟踪方法不需要进行在线学习、在线分类及特征融合等相关操作,通过视觉量子熵能量计算、频率阶跃不变性提取、量子平衡跟踪以及目标状态预测这 4 个步骤实行目标跟踪,在时间鲁棒性和空间鲁棒性上凸显出较强的优势.同时,由图 26 可知,Benchmark 性能评估体系提供测试的视频数据在目标运动速度、形状变化幅度、尺度伸缩强度、相似背景/目标遮挡与干扰、光照强度变化、特征域旋转、系统与噪声、杂波等多种复杂因素的综合影响下,对目标跟踪方法的测试具有全面性和客观性,而且其评估过程严格,视频数据环境贴近实际现场,评估指标较为公正.因此,本文提出的视觉量子跟踪方法在跟踪的准确性、稳定性、实时性及鲁棒性上都显示出了较强的优势和很好的跟踪性能.同时,由于视觉量子跟踪方法在进行量子频率阶跃计算、量子平衡状态移动过程中对图像采集设备的采集频率具有较强的依赖性,较高的图像采集频率对本文视觉量子目标跟踪的准确性和目标状态预测准确性具有积极的推动作用,因此,如果能够在提升图像采集设备的采集频率前提下进行视觉量子跟踪,将会使本文方法发挥出更好的跟踪性能.

### 3.4 计算复杂性与数据量分析

在 Pentium4 3.06G CPU/1G 内存的机器上,对一组大小为  $a \times b$ 、长度为  $l$  的视频进行目标跟踪,手动选取的目标跟踪窗大小为  $m \times n$ ,视觉量子数量为  $N$ .

本文方法的时间计算量包含 3 个部分:第 1 部分是量子频率计算时间,第 2 部分是量子平衡跟踪的计算时间,第 3 部分是目标状态预测的计算时间.计算量子频率信息的时间复杂度为  $O(M \log N)$ ,进行量子平衡跟踪的时间复杂度为  $O(M \log N + N)$ ,进行目标状态预测的时间复杂度为  $O[(mn \cdot \log N)/N]$ .因此,本文方法的时间复杂度是  $O[(2N + mn/N) \log N + N]$ ,远小于基于 TLD 的时间复杂度  $O(abm^3 n^3 N^2)$ ,也远小于基于 OLSL 的时间复杂度  $O(mnN^2 + \log N)$  和基于 WLMC 的时间复杂度  $O(m^2 n^2 N^2 + mnN)$ .本文方法的空间复杂度为  $O(3N)$ ,即  $O(N)$ ,而文献[31-33]方法的空间复杂度都是  $O(N^2)$ ,因此,本文方法的计算效率优于 TLD 方法,也优于 OLSL 和 WLMC 的方法.

在数据处理方面,本文方法需要处理的数据主要集中在目标混合边缘区域的量子频率计算上,数据量较小;同时,采用极大似然估计方法预测目标位置,无需对候选目标进行全局搜索和存储多种目标姿态模型,降低了目标搜索范围和待处理的数据总量.对于本节设定的视频图像,其处理的数据总量远远小于  $abl$  个像素,而文献[31-33]方法需要处理的数据总量均为  $abl$  个像素.因此,本文方法处理的数据量小于其他 3 种方法的数据量.

## 4 结论及展望

本文在生物“视觉单元”协同工作的方向上提出了一种新的基于视觉量子的目标跟踪方法.视觉量子目标跟踪方法是建立在子空间目标协同分析的整体框架上,利用视觉量子采样、熵能量计算、频率阶跃不变性提取、量子平衡跟踪以及采用极大似然估计预测运动目标状态,重复验证移动后的视觉量子的量子平衡状态,确保目标跟踪有效性.与现有目标跟踪方法相比,该方法的优势在于:

- (1) 抓住了变结构运动目标前景与背景交界处具有的能量频率阶跃不变性的特点;
- (2) 将阶跃不变特征采用多个具有独立性和约束性的视觉量子进行描述,可以有效消除形状变化、尺度变化等变结构因素对运动目标跟踪的影响,失跟率较低,鲁棒性能较高;
- (3) 视觉量子数据量较小,同时,不需要存储目标多种姿态模型,计算复杂度和空间复杂度较低,跟踪实时性较高.

通过 IBM 评估测试和 Benchmark 评估测试表明,视觉量子在多种复杂环境下,特别是速度较高的变结构运动目标跟踪中,保持了较强的稳定性,对环境变化和自身变化具有很好的鲁棒性和准确性,准确率平均为 91.5%,在高速变结构机动目标跟踪、IBM 评估数据库和 Benchmark 评估数据库上跟踪的实时性平均为 72 帧/秒.本文方法不仅对视频目标跟踪有效,而且对三维场景图像检索、视觉场景分析也具有普遍意义,为计算机视觉处理提供了新的思路.

视觉量子通过目标子空间量子簇的协同分析进行目标跟踪,当目标图像整体模糊程度较高(熵分布趋于 1)时,量子频率近似程度较为严重,影响了目标跟踪效果.虽然模糊程度较高的视频图像在视觉意义上已经失去作

用,但针对模糊图像跟踪做进一步分析,将对视觉量子跟踪性能的完善起到很好的推动作用.因此,本文今后将针对模糊图像的目标跟踪做进一步的研究工作.

**致谢** 在此,我们向对本文的工作给予支持、指导和建议的国内外同行,尤其是英国 Surrey 大学学者 Zdenek Kalal、韩国 Seoul 大学学者 Junseok Kwon、美国 Carnegie Mellon 大学的 Robert T. Collins 教授表示感谢.

#### References:

- [1] David T, Matthew F, Atsushi N. Motion coherent tracking using multi-label MRF optimization. *Int'l Journal of Computer Vision*, 2012,100(2):190–202. [doi: 10.1007/s11263-011-0512-5]
- [2] Xi L, Dick A, Chun HS, Van DHA, Han ZW. Incremental learning of 3D-DCT compact representations for robust visual tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2013,35(4):863–881. [doi: 10.1109/TPAMI.2012.166]
- [3] Salmond DJ, Gordon NJ. Group and extended object tracking. In: *Proc. of IEE Colloquium on Target Tracking*. London: IEEE SPIE, 1999. 1611–1614. [doi: 10.1049/ic:19990517]
- [4] Waxmann MJ, Drummond OE. A bibliography of cluster (group) tracking. In: *Proc. of the Int'l Conf. on Signal and Data Processing of Small Targets*. Orlando: SPIE, 2004. 551–560. [doi: 10.1117/12.548357]
- [5] Bugeau A, Perez P. Detection and segmentation of moving objects in highly dynamic scenes. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. Minnesota: IEEE, 2007. 1–8. [doi: 10.1109/CVPR.2007.383244]
- [6] Lu Z, Laurens VM. Structure preserving object tracking. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. Piscataway: IEEE Press, 2013. 1838–1845. [doi: 10.1109/CVPR.2013.240]
- [7] Rizwan C, Gregory H, Rene V. Dynamic template tracking and recognition. *Int'l Journal of Computer Vision*, 2013,105(1):19–48. [doi: 10.1007/s11263-013-0625-0]
- [8] Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 1993,15(9):850–863. [doi: 10.1109/34.232073]
- [9] Huttenlocher DP, Noh JJ, Rucklidge WJ. Tracking nonrigid objects in complex scenes. In: *Proc. of the 4th Int'l Conf. on Computer Vision*. Piscataway: IEEE Press, 1993. [doi: 10.1109/ICCV.1993.378231]
- [10] Haritaoglu I, Harwood D, Davis LS. W<sup>4</sup>: Real-Time surveillance of people and their activities. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2008,22(8):809–830. [doi: 10.1109/34.868683]
- [11] Stauffer C, Grimson WEL. Learning patterns of activity using real-time tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2006,22(8):747–757. [doi: 10.1109/34.868677]
- [12] Rezaeian M, Vo BN. Error bounds for joint detection and estimation of a single object with random finite set observation. *IEEE Trans. on Signal Processing*, 2010,58(3):1493–1506. [doi: 10.1109/TSP.2009.2037665]
- [13] Yan XX, Han CZ. Multiple target tracking based on online estimation of clutter intensity. *Control and Decision*, 2012,27(4):507–512 (in Chinese with English abstract). [doi: 10.13195/j.cd.2012.04.29.yanxx.013]
- [14] Juan R, Raul E, Victor A. A fuzzy inference approach to template-based visual tracking. *Machine Vision and Applications*, 2012, 23(3):427–439. [doi: 10.1007/s00138-010-0314-8]
- [15] Airouche M, Bentabet L, Zelmat M, Gao G. Pedestrian tracking using color, thermal and location cue measurements: A DSMT-based framework. *Machine Vision and Applications*, 2012,23(5):999–1010. [doi: 10.1007/s00138-011-0342-z]
- [16] Zhou HY, Yuan Y, Shi CM. Object tracking using SIFT features and mean shift. *Computer Vision and Image Understanding*, 2009, 113(3):345–352. [doi: 10.1016/j.cviu.2008.08.006]
- [17] Maggio E, Cavallaro A. Learning scene context for multiple object tracking. *IEEE Trans. on Image Processing*, 2009,18(8):1873–1884. [doi: 10.1109/TIP.2009.2019934]
- [18] Carl YR, Victor P, Ian R. Regressing local to global shape properties for online segmentation and tracking. *Int'l Journal of Computer Vision*, 2013,106(3):269–281. [doi: 10.1007/s11263-013-0635-y]
- [19] Vo BN, Ma WK. The Gaussian mixture probability hypothesis density filter. *IEEE Trans. on Signal Processing*, 2009,51(11):4091–4104. [doi: 10.1109/TSP.2006.881190]

- [20] Mahler R. Detecting, tracking, and classifying group targets: A unified approach. In: Proc. of the Int'l Conf. on Signal and Data Processing of Small Targets. Orlando: SPIE, 2009. 217–228. [doi: 10.1117/12.436950]
- [21] Xi L, Wei MH, Chun HS, Zhong FZ, Anthony D, Anton VDH. A survey of appearance models in visual object tracking. *ACM Trans. on Intelligent Systems and Technology*, 2013,4(4):478–488. [doi: 10.1145/2508037.2508039]
- [22] Comaniciu D, Ramesh V, Meer P. Real-Time tracking of non-rigid objects using mean shift. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2000. 142–149. [doi: 10.1109/CVPR.2000.854761]
- [23] Comaniciu D, Ramesh V, Meer P. Kernel-Based object tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2007, 25(5):564–577. [doi: 10.1109/TPAMI.2003.1195991]
- [24] Yang C, Duraiswami R, Davis L. Efficient means shift tracking via a new similarity measure. In: Proc. of the 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2005. 176–183. [doi: 10.1109/CVPR.2005.139]
- [25] Jeong MH, You BJ, Oh YH, Oh SR, Oh SH. Adaptive mean-shift tracking with novel color model. In: Proc. of the IEEE Int'l Conf. on Mechatronics & Automation. Niagara Falls, 2005. 1329–1333. [doi: 10.1109/ICMA.2005.1626746]
- [26] Davy M, Tourneret JY. Generative supervised classification using dirichlet process priors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2010,32(10):1781–1794. [doi: 10.1109/TPAMI.2010.21]
- [27] Liam E, Nicholas D, Jiri M, Richard B. Linear regression and adaptive appearance models for fast simultaneous modeling and tracking. *Int'l Journal of Computer Vision*, 2011,95(2):154–179. [doi: 10.1007/s11263-010-0364-4]
- [28] Ying LT, Andrew S, Max L. Robust and efficient foreground analysis in complex surveillance videos. *Machine Vision and Applications*, 2012,23(5):967–983. [doi: 10.1007/s00138-011-0377-1]
- [29] Yi W, Jongwoo L, Hsuan Y. Online object tracking: A benchmark. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2013. 2411–2418. [doi: 10.1109/CVPR.2013.312]
- [30] Lü XB, Zhou QB, Chen ZM, Xiong YY, Cai K. The Gaussian mixture probability hypothesis density filter and its application to multi-target tracking. *Chinese Journal of Computers*, 2012,35(2):397–403 (in Chinese with English abstract). [doi: 10.3724/sP.J.1016.2012.00397]
- [31] Kalal Z, Mikolajczyk K, Matas J. Tracking-Learning-Detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2012, 34(7):1409–1422. [doi: 10.1109/TPAMI.2011.239]
- [32] Rui Y, Qinfeng S, Chunhua S, Yanning Z, Vandenhengel A. Part-Based visual tracking with online latent structural learning. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2013. 2363–2370. [doi: 10.1109/CVPR.2013.306]
- [33] Junseok K, Kyoung ML. Wang-Landau Monte Carlo-based tracking methods for abrupt motions. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2013,35(4):1011–1024. [doi: 10.1109/TPAMI.2012.161]
- [34] Hu DK. The research on technology of image understanding based on biological visual perception [Ph.D. Thesis]. Chengdu: University of Electronic Science and Technology of China, 2012 (in Chinese with English abstract).
- [35] Ali B, Laurent I. State-of-the-Art in visual attention modeling. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2013, 35(1):185–207. [doi: 10.1109/TPAMI.2012.89]
- [36] Zhang JS, Xu ZB, Ruan YW. A new data processing method based on a biological model of compound eye: Direction quantization representation and its theory. *Chinese Journal of Computers*, 2000,23(6):609–614 (in Chinese with English abstract). [doi: 10.3321/j.issn:0254-4164.2000.06.007]
- [37] Zhang JS, Xu ZB, Ruan YW. A new data processing method based on a biological model of compound eye: Direction quantization representation in applications. *Chinese Journal of Computers*, 2000,23(6):615–619 (in Chinese with English abstract). [doi: 10.3321/j.issn:0254-4164.2000.06.008]
- [38] Goodman N. Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, 1963,34(1):152–177. [doi: 10.1214/aoms/1177704250]
- [39] Brown LM, Senior AW, Tian YL, Connell J, Hampapur A. Performance evaluation of surveillance systems under varying conditions. In: Proc. of the IEEE Int'l Workshop on Performance Evaluation of Tracking and Surveillance. Piscataway: IEEE Press, 2005. 1–8.

- [40] Grabner H, Roth PM, Bischof H. Is pedestrian detection really a hard task. In: Proc. of the 10th IEEE Int'l Workshop on Performance Evaluation of Tracking and Surveillance. Piscataway: IEEE Press, 2007. 1–8.

**附中文参考文献:**

- [13] 闫小喜,韩崇昭.基于杂波强度在线估计的多目标跟踪算法.控制与决策,2012,27(4):507–512. [doi: 10.13195/j.cd.2012.04.29.yanxx.013]
- [30] 吕学斌,周群彪,陈正茂,熊运余,蔡葵.高斯混合概率假设密度滤波器在多目标跟踪中的应用.计算机学报,2012,35(2):397–403. [doi: 10.3724/sP.J.1016.2012.00397]
- [34] 胡德昆.基于生物视觉感知机制的图像理解技术研究[博士学位论文].成都:电子科技大学,2012.
- [36] 张讲社,徐宗本,阮奕文.一类模拟生物复眼几何结构的数据处理方法——方向量子化表示的理论.计算机学报,2000,23(6): 609–614. [doi: 10.3321/j.issn:0254-4164.2000.06.007]
- [37] 张讲社,徐宗本,阮奕文.一类模拟生物复眼几何结构的数据处理方法——方向量子化表示的应用.计算机学报,2000,23(6): 615–619. [doi: 10.3321/j.issn:0254-4164.2000.06.008]



姜文涛(1986—),男,辽宁大连人,博士,副教授,CCF 专业会员,主要研究领域为图像与视觉信息计算,视频图像理解.



袁姮(1989—),女,博士生,主要研究领域为信息管理与信息系统,人工智能.



刘万军(1959—),男,教授,博士生导师,CCF 高级会员,主要研究领域为计算机视觉图像理解,模式识别.



张海涛(1974—),男,博士,教授,CCF 专业会员,主要研究领域为计算机视觉,图形图像处理.