基于中智加权相似度量的尺度自适应视觉目标跟踪算法

胡珂立[1]，范恩[1]，叶军[1]，沈士根[1]，谷宇章[2]
（1. 绍兴文理学院，浙江 绍兴 312000;
2. 中国科学院上海微系统与信息技术研究所，上海 200050）

摘　要：中智集理论是传统模糊理论的拓展，它能够表述现实生活中的不确定性信息。面对不同问题，中智框架下的真（truth）、不确定（indeterminacy）、假（falsity）隶属度所占权重可能不同。提出了一种分量加权的余弦相似度量，并将其引入均值漂移视觉跟踪算法中。首先基于 3σ 理论和目标/背景相似度两种属性提出了相应的真、不确定、假量测，然后利用加权余弦相似度量构建权值向量，同时提出了基于中智加权余弦相似度量的尺度更新算法，综合提升均值漂移跟踪性能。实验结果表明，提出的视觉跟踪算法能较好克服相似背景、光照变化、尺度变化等挑战。

关键词：中智集；加权相似度量；目标跟踪
中图分类号：TP391
文献标识码：A
doi: 10.11959/j.issn.1000−0801.2018176

A scale adaptive visual object tracking algorithm based on weighted neutrosophic similarity coefficient

HU Keli[1], FAN En[1], YE Jun[1], SHEN Shigen[1], GU Yuzhang[2]
1. Shaoxing University, Shaoxing 312000, China
2. Shanghai Institute of Microsystem and Information Technology,
Chinese Academy of Sciences, Shanghai 200050, China

Abstract: The weight of the truth, indeterminacy, and falsity membership under the neutrosophic framework may be different when dealing with different problems. Due to this, a component weighted cosine similarity coefficient was proposed, and it was introduced into the mean shift tracking algorithm. Firstly, the corresponding methods for calculating the membership of the truth, indeterminacy, and falsity were proposed based on the theory of 3σ, as well as the similarity between the features of the corresponding area of the object and background. Then the weighted cosine similarity coefficient was used to construct the weight vector. In addition, a weighted cosine similarity coefficient based scale updating method was proposed. The experimental results demonstrate that the modified visual tracking algorithm performs well, even when there exists challenges like similar background, illumination or scale variation.

Key words: neutrosophic set, weighted similarity coefficient, object tracking

收稿日期：2017−12−11；**修回日期**：2018−05−07
通信作者：胡珂立，ancimoon@gmail.com
基金项目：国家自然科学基金资助项目（No.61603258, No.61703280）；浙江省公益技术应用研究项目（No.2016C31082）
Foundation Items：The National Natural Science Foundation of China (No.61603258, No.61703280), The Public Welfare Technology Application Research Project of Zhejiang Province (No.2016C31082)

1　引言

随着“平安城市”的进一步建设以及人们对安全生活、生产需求的进一步提升，以物联网为代表的多项技术不断发展。视觉传感器是物联网感知世界的关键途径，众多视觉传感器产生的海量数据须经编码[1]传输，与视觉传感数据紧密相关的视觉分析[2,3]也是当前最重要的环境感知手段之一。视觉目标跟踪[4-6]的任务为锁定监控区域内兴趣目标，时刻感知其位置信息。视觉目标跟踪是进行后续高级分析任务的重要基础。由于现实场景中面临诸多挑战，提升视觉目标跟踪性能仍是一个挑战性问题。

均值漂移算法由 Comaniciu 等人[7, 8]提出，并被用于视觉目标跟踪。采用颜色直方图作为目标特征。在目标跟踪过程中，算法通过最小化跟踪目标与候选目标区域概率密度函数的距离来确定当前目标位置。由于均值漂移跟踪算法采用颜色直方图特征和高效的目标区域定位策略，能有效地克服运动模糊、目标形变等挑战。然而，当背景区域与目标特征相近时，该算法极易发生偏移，最终导致跟踪失败。鉴于此种问题，一些新的度量方式或特征被引入，例如 Cross-Bin 度量[9]、SIFT 特征[10]、纹理特征[11]，在一定程度上提升了算法顽健性。Tomas 等人[12]将背景信息引入，并提出了基于背景加权的方法。在目标尺度估计方面，Tomas 等人[12]采用了前向后向稳定性检测方法，经改进的此均值漂移算法的表现超过了多数性能优异的跟踪算法。此外，在尺度更新方面，Robert 等人[13]采用差分尺度空间滤波器的策略，也取得了一定成效。尽管众多学者做了很多努力，但视觉目标跟踪仍是一个开放问题。鉴于均值漂移跟踪算法的高效性，寻求一种恰当的方式以充分利用目标信息和背景信息提升目标位置估计的顽健性仍十分必要。

中智集理论[14]由 Smarandache 提出，相较于传统的模糊理论，中智集理论除了对真（truth）、假（falsity）分量进行表述外，还专门对不确定（indeterminacy）分量进行了处理。中智集理论在处理不确定信息方面展现出很大优势。为了将中智集理论用于工程领域，参考文献[15]提出了单值中智集的概念，将真、假、不确定分量的值限定在 0 到 1 的闭区间范围内。为将单值中智集用于工程决策，参考文献[16]提出了余弦相似量测。鉴于中智理论的有效性，其被广泛用于计算机视觉[17-23]、汽轮机故障诊断[24]、治疗方案选择[25,26]、聚类理论等[17,27-29]。参考文献[17,18,21-23]将中智集理论用于图像分割，其中参考文献[17]提出了中智相似聚类方法；参考文献[18]将不确定滤波器引入图像分割；参考文献[21]引入核约束，提出了核空间中智模糊均值聚类算法。在参考文献[19]中，率先将单值中智集引入视频目标跟踪算法中，提出了面向图像颜色域和深度域的中智真、假、不确定分量的度量方法，并利用单值中智交叉熵完成了颜色和深度空间的信息融合，以此提升了跟踪算法性能。参考文献[30]利用中智相似度量建立新的权值，一定程度上提升了跟踪算法的顽健性，然而其将 3 个中智分量同等看待，且采用固定尺度信息，一定程度上限制了算法性能。

本文引入中智加权相似度量，提出了尺度自适应的中智均值漂移视觉跟踪算法，主要贡献如下：

（1）提出了中智分量加权的思路，并将其引入余弦相似度量；

（2）针对帧间目标特征稳定度和目标背景差异性提出了相应属性下的真、不确定、假量测计算方法，并利用加权余弦相似度量建立权值向量；

（3）针对目标尺度自适应更新问题提出了目标尺度缩小/扩大的真、不确定、假量测计算方法，

并提出了中智框架下的尺度更新策略。

2 中智余弦相似度量

2.1 单值中智集

单值中智集是中智集的重要组成部分，单值中智集的提出使得中智理论用于科学与工程问题成为可能。单值中智集中每个元素对应的真、不确定和假分量均为实数数值。假定 X 是一个全集，在 X 中的单值中智集 A 可以表示为：

$$A=\left\{\left\langle x,T_A(x),I_A(x),F_A(x)\right\rangle \middle| x\in X\right\} \tag{1}$$

其中，$T(x)$、$I(x)$和 $F(x)$分别表示对应元素的真、不确定和假分量量测函数，且 $T_A(x)$、$I_A(x)$、$F_A(x)\in[0,1]$，$0\leqslant T_A(x)+I_A(x)+F_A(x)\leqslant 3$。

假定 $A=\{A_1,A_2,\cdots,A_m\}$ 是一个可供选择的集合，$C=\{C_1,C_2,\cdots,C_n\}$ 是一个条件属性集合，依照式(1)，A_i 可以被表示为：

$$A_i=\left\{\left\langle C_j,T_{C_j}(A_i),I_{C_j}(A_i),F_{C_j}(A_i)\right\rangle \middle| C_j\in C\right\} \tag{2}$$
$$i=1,2,\cdots,m,j=1,2,\cdots,n$$

其中，$T_{Cj}(A_i)$、$I_{Cj}(A_i)$和 $F_{Cj}(A_i)$是在条件属性 C_j 下针对 A_i 的真、不确定和假分量量测函数，且 $T_{Cj}(A_i)$、$I_{Cj}(A_i)$、$F_{Cj}(A_i)\in[0,1]$，$0\leqslant T_{Cj}(A_i)+I_{Cj}(A_i)+F_{Cj}(A_i)\leqslant 3$。多条件属性下的集合 A 为多属性单值中智集。

2.2 余弦相似度量

参考文献[31]中提出了一种面向多属性单值中智集的余弦（向量形式）相似度量测方法，A_i 与 A_j 在属性 C_k 下的相似度可以表示为：

$$S_{C_k}(A_i,A_j)=\frac{T_{C_k}(A_i)T_{C_k}(A_j)+I_{C_k}(A_i)I_{C_k}(A_j)+F_{C_k}(A_i)F_{C_k}(A_j)}{\sqrt{T_{C_k}^2(A_i)+I_{C_k}^2(A_i)+F_{C_k}^2(A_i)}\sqrt{T_{C_k}^2(A_j)+I_{C_k}^2(A_j)+F_{C_k}^2(A_j)}} \tag{3}$$

假定 A^*代表中智集范畴内理想情况下的选择，则任一选择 A_i 同 A^*的加权相似度可以表示为：

$$W(A_i,A^*)=\sum_{k=1}^{n}w_k\frac{T_{C_k}(A_i)T_{C_k}(A^*)+I_{C_k}(A_i)I_{C_k}(A^*)+F_{C_k}(A_i)F_{C_k}(A^*)}{\sqrt{T_{C_k}^2(A_i)+I_{C_k}^2(A_i)+F_{C_k}^2(A_i)}\sqrt{T_{C_k}^2(A^*)+I_{C_k}^2(A^*)+F_{C_k}^2(A^*)}} \tag{4}$$

其中，$w_k\in[0,1]$，w_k 是对应于属性 C_k 的权值，且 $\sum_{k=1}^{n}w_k=1$。

2.3 加权余弦相似度量

观察式（4），可以发现该加权相似度仅是对不同属性的加权，然而在实际应用中，有时须强化各分量内部的 T、I、F 分量中的一个或多个值。现有的中智相似度量算法并没有这项考虑，因此，将 A_i 同 A_j 之间的相似度表示为：

$$W'(A_i,A_j)=\sum_{k=1}^{n}w_k\frac{\alpha T_{C_k}(A_i)T_{C_k}(A_j)+\beta I_{C_k}(A_i)I_{C_k}(A_j)+\gamma F_{C_k}(A_i)F_{C_k}(A_j)}{\sqrt{\alpha T_{C_k}^2(A_i)+\beta I_{C_k}^2(A_i)+\gamma F_{C_k}^2(A_i)}\sqrt{\alpha T_{C_k}^2(A_j)+\beta I_{C_k}^2(A_j)+\gamma F_{C_k}^2(A_j)}} \tag{5}$$

其中，α、β、$\gamma\in[0,1]$，分别为 T、I、F 分量的权值，且满足 $\alpha+\beta+\gamma=1$。可以看到，式（5）中的主体部分实际是计算向量 $(\sqrt{\alpha}T_{C_k}(A_i),\sqrt{\beta}I_{C_k}(A_i),\sqrt{\gamma}F_{C_k}(A_i))$ 和向量 $(\sqrt{\alpha}T_{C_k}(A^*),\sqrt{\beta}I_{C_k}(A^*),\sqrt{\gamma}F_{C_k}(A^*))$ 之间的余弦值。因此，很容易得出如下 3 个结论：

（P1） $0\leqslant W'(A_i,A_j)\leqslant 1$；

（P2） $W'(A_i,A_j)=W'(A_j,A_i)$；

（P3） $W'(A_i,A_j)=1$，如果 $A_i=A_j$。

假定 A^*代表理想选择，则 A_i 同 A^*之间的相似度可表示为：

$$W'(A_i,A^*)=\sum_{k=1}^{n}w_k\frac{\alpha T_{C_k}(A_i)T_{C_k}(A^*)+\beta I_{C_k}(A_i)I_{C_k}(A^*)+\gamma F_{C_k}(A_i)F_{C_k}(A^*)}{\sqrt{\alpha T_{C_k}^2(A_i)+\beta I_{C_k}^2(A_i)+\gamma F_{C_k}^2(A_i)}\sqrt{\alpha T_{C_k}^2(A^*)+\beta I_{C_k}^2(A^*)+\gamma F_{C_k}^2(A^*)}} \tag{6}$$

3 权值向量构建

为提升均值漂移跟踪算法的顽健性，本文以

目标特征、背景特征为出发点，从帧间稳定度和帧内相似度两个方面对均值漂移跟踪算法进行约束。

3.1　均值漂移跟踪算法

均值漂移算法通常采用核加权颜色直方图，在选定跟踪目标后（通常用矩形框选定目标区域），目标特征被表示为：

$$\hat{q}_u = C\sum_{i=1}^{n} k\left(\left\|\boldsymbol{x}_i^*\right\|^2\right)\delta\left[b\left(\boldsymbol{x}_i^*\right)-u\right] \tag{7}$$

其中，$\hat{q}_u$ 是直方图 $\hat{q}$ 的一个分量，假定 $\hat{q}$ 有 m 个分量，则有 $\sum_{u=1}^{m}\hat{q}_u = 1$。$\boldsymbol{x}_i^*$ 为相对矩形框中心的图像坐标，$b(x)$是一个映射函数，它将 x 像素点位置的颜色信息映射在 1～m 区间内的某一个数值上。$k(x)$为核函数，核函数要求具有等向性、凸和单调递减性。$\delta(x)$是克罗内克函数，C 为归一化参数，n 为目标区域像素点总数。

在跟踪过程中，需要选取候选目标，提取候选目标直方图特征与目标特征比对。假定 y 是候选目标区域（同尺度矩形框）的中心坐标，候选目标特征可表示为：

$$\hat{p}_u = C_h\sum_{i=1}^{n_h} k\left(\left\|\frac{\boldsymbol{y}-\boldsymbol{x}_i}{h}\right\|^2\right)\delta\left[b\left(\boldsymbol{x}_i-u\right)\right] \tag{8}$$

其中，x_i 为候选区域内的像素点坐标，n_h 为候选区域总像素点数，C_h 为归一化参数。

巴氏相关系数被用来衡量目标特征和候选目标特征的相似度，该相似度被表示为：

$$\rho\left[\boldsymbol{p}(\hat{\boldsymbol{y}}),\hat{\boldsymbol{q}}\right] = \sum_{u=1}^{m}\sqrt{\hat{p}_u(\hat{\boldsymbol{y}})\hat{q}_u} \tag{9}$$

上一帧确定的目标区域通常会被作为下一帧的首个候选目标区域。新一帧目标区域须使得 $\rho\left[\boldsymbol{p}(\hat{\boldsymbol{y}}),\hat{\boldsymbol{q}}\right]$ 最大。针对该项目标，假设当前候选目标区域中心点图像坐标为 $\hat{\boldsymbol{y}}_0$，则下一个候选目标中心点坐标应为[7,8]：

$$\hat{\boldsymbol{y}}_1 = \frac{\sum_{i=1}^{n_h}\boldsymbol{x}_i w_i g\left(\left\|\frac{\hat{\boldsymbol{y}}_0-\boldsymbol{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{\hat{\boldsymbol{y}}_0-\boldsymbol{x}_i}{h}\right\|^2\right)} \tag{10}$$

其中，$g(x) = -k'(x)$，式（10）中的 w_i 可通过式（11）计算：

$$w_i = \sum_{u=1}^{m}\delta\left[b\left(\boldsymbol{x}_i\right)-u\right]\sqrt{\frac{\hat{q}_u}{\hat{p}_u\left(\hat{y}_0\right)}} \tag{11}$$

3.2　多属性中智度量

为顽健完成视觉目标跟踪问题，须强化帧间特征较稳定的特征以及和背景特征差异较大的特征。鉴于此，本文以直方图特征每个分量为分析对象，提出了两种基于单值中智集的 T、I、F 量测。

对于帧间稳定条件属性 C_s，借助 3σ 理论，t 时刻的 T、I、F 量测分别被定义为：

$$T_{C_s}^t(u) = \frac{1}{N}\mathrm{Num}(\left|q_u^i-\mu\right| \leqslant \sigma, q_u^i \in \boldsymbol{Q}_u^N) \tag{12}$$

$$I_{C_s}^t(u) = \frac{1}{N}\mathrm{Num}(\sigma < \left|q_u^i-\mu\right| \leqslant 2\sigma, q_u^i \in \boldsymbol{Q}_u^N) \tag{13}$$

$$F_{C_s}^t(u) = \frac{1}{N}\mathrm{Num}(\left|q_u^i-\mu\right| > 2\sigma, q_u^i \in \boldsymbol{Q}_u^N) \tag{14}$$

其中，$\boldsymbol{Q}_u^N$ 为 t–N+1 时刻到 t 时刻目标跟踪位置特征直方图中的第 u 个分量的集合，μ 为该集合的均值，σ 为该集合的标准差，N 表示该集合中总元素数。可以发现，$T_{C_s}^t(u)$、$I_{C_s}^t(u)$ 和 $F_{C_s}^t(u)$ 分量的含义分别为 t 时刻前连续 N 帧时序内满足条件 $\left|q_u^i-\mu\right| \leqslant \sigma$、$\sigma < \left|q_u^i-\mu\right| \leqslant 2\sigma$ 和 $\left|q_u^i-\mu\right| > 2\sigma$ 的 u 分量数量占总元素数的概率。

对于目标特征强度及目标/背景差异属性 C_b，相应的 t 时刻的 T、I、F 量测分别被定义为：

$$T_{C_b}^t(u) = \hat{q}_u \tag{15}$$

$$I_{C_b}^t(u) = b_u \tag{16}$$

$$F_{C_b}^t(u) = 1 - T_{C_b}^t(u) \tag{17}$$

其中，$\hat{q}_u$ 为目标特征直方图的第 u 个分量，b_u 为

背景区域特征直方图的第 u 个分量。在本文中，背景区域被定义为目标区域的等比例外扩区域，两个区域中心均为原目标区域中心。假定目标区域为 G_o，则背景区域 $G_b=\lambda G_o-G_o$，λ 为扩大系数。b_u 的计算依据式（7），本文并非在每一帧均更新 b_u，而是当前帧背景区域特征与 b_u 差异较大时才将其作为新的背景特征。

3.3 权值计算

将式（12）~式（17）代入式（6），便可以得到直方图中第 u 个分量的中智权值：

$$\begin{aligned} w_u^{NSt} &= w_{C_s} W(u_{cs}, A^*) + w_{C_b} W(u_{cb}, A^*) \\ &= w_{C_s} \frac{\alpha T_{C_s}^t(u) T_{C_s}^t(A^*) + \beta I_{C_s}^t(u) I_{C_s}^t(A^*) + \gamma F_{C_s}^t(u) F_{C_s}^t(A^*)}{\sqrt{\alpha T_{C_s}^{t\,2}(u) + \beta I_{C_s}^{t\,2}(u) + \gamma F_{C_s}^{t\,2}(u)} \sqrt{\alpha T_{C_s}^{t\,2}(A^*) + \beta I_{C_s}^{t\,2}(A^*) + \gamma F_{C_s}^{t\,2}(A^*)}} \\ &+ w_{C_b} \frac{\alpha T_{C_b}^t(u) T_{C_b}^t(A^*) + \beta I_{C_b}^t(u) I_{C_b}^t(A^*) + \gamma F_{C_b}^t(u) F_{C_b}^t(A^*)}{\sqrt{\alpha T_{C_b}^{t\,2}(u) + \beta I_{C_b}^{t\,2}(u) + \gamma F_{C_b}^{t\,2}(u)} \sqrt{\alpha T_{C_b}^{t\,2}(A^*) + \beta I_{C_b}^{t\,2}(A^*) + \gamma F_{C_b}^{t\,2}(A^*)}} \end{aligned} \tag{18}$$

其中，w_{C_s}、w_{C_b} 分别为两个条件属性对应的权值，w_{C_s}、$w_{C_b} \in [0,1]$，且 $w_{C_s}+w_{C_b}=1$。此处的理想选择 A^* 在任一属性条件下均被定义为 $A^*=\langle 1,0,0 \rangle$，由此，式（18）可以简化为：

$$\begin{aligned} w_u^{NSt} &= w_{C_s} \frac{\sqrt{\alpha} T_{C_s}^t(u)}{\sqrt{\alpha T_{C_s}^{t\,2}(u) + \beta I_{C_s}^{t\,2}(u) + \gamma F_{C_s}^{t\,2}(u)}} + \\ & w_{C_b} \frac{\sqrt{\alpha} T_{C_b}^t(u)}{\sqrt{\alpha T_{C_b}^{t\,2}(u) + \beta I_{C_b}^{t\,2}(u) + \gamma F_{C_b}^{t\,2}(u)}} \end{aligned} \tag{19}$$

4 尺度自适应方法

在跟踪目标的过程中，目标成像大小可能会发生变化，当跟踪目标框的尺度变化长时间无法适应目标成像尺寸时，易发生跟踪失败。本文将目标候选区域特征与目标特征的相似度作为可靠信息，将目标候选区域特征与背景区域特征的相似度作为不确定信息，首次将中智集理论引入目标尺度更新策略中。

在当前帧目标位置确定后，开始分析是否需要进行尺度更新。首先考虑尺度缩小的情况，基于本文对可靠信息和不确定信息的分析，缩小尺度对应的 T、I、F 量测分别被定义为：

$$T_{\mathrm{sa}} = \rho\left[\boldsymbol{p}_{\mathrm{sa}}(\hat{\boldsymbol{y}}), \hat{\boldsymbol{q}}\right] \tag{20}$$

$$I_{\mathrm{sa}} = \rho\left[\boldsymbol{p}_{\mathrm{sa}}(\hat{\boldsymbol{y}}), \boldsymbol{b}_{\mathrm{sa}}(\hat{\boldsymbol{y}})\right] \tag{21}$$

$$F_{\mathrm{sa}}(u) = 1 - T_{\mathrm{sa}} \tag{22}$$

其中，$\boldsymbol{p}_{\mathrm{sa}}(\hat{\boldsymbol{y}})$ 为当前帧目标区域缩小后对应小区域的特征直方图，确定缩小区域的方式与式（16）中背景区域的确定方法类似，同样以当前目标区域为参考，区域中心不变，长宽按同比例缩小 λ_{sa} 倍，λ_{sa} 为对应的缩放系数。$\hat{\boldsymbol{q}}$ 为初始目标直方图。$\boldsymbol{b}_{\mathrm{sa}}(\hat{\boldsymbol{y}})$ 为当前区域对应背景区域的特征直方图。容易发现，T_{sa} 为目标区域缩小后与初始目标特征的相似度，I_{sa} 为目标区域缩小后对应背景区域与 $\boldsymbol{p}_{\mathrm{sa}}(\hat{\boldsymbol{y}})$ 的相似度。

与式（18）类似，此处的理想选择 A^* 在任一属性条件下也被定义为 $A^*=\langle 1,0,0 \rangle$，将式（20）~式（22）代入式（6），可以得到：

$$w_{\mathrm{sscal}} = \frac{\sqrt{\alpha} T_{\mathrm{sa}}}{\sqrt{\alpha T_{\mathrm{sa}}^{\ 2} + \beta I_{\mathrm{sa}}^{\ 2} + \gamma F_{\mathrm{sa}}^{\ 2}}} \tag{23}$$

依据同样的策略，可以获得放大尺度对应的中智相似度权值 w_{bscal}，其中放大系数为 λ_{ba}。本文取 $\lambda_{\mathrm{sa}} = \lambda_{\mathrm{ba}}$。

得到 w_{sscal} 和 w_{bscal} 后，可依照式（24）确定当前帧尺度：

$$\lambda_{\mathrm{new}} = \begin{cases} 1-\lambda_0, w_{\mathrm{sscal}} > s w_{\mathrm{bscal}} \\ 1+\lambda_0, w_{\mathrm{bscal}} > s w_{\mathrm{sscal}} \\ 1, \text{其他} \end{cases} \tag{24}$$

其中，s 为比例因子，且满足 $s>1$。该比例因子的设定是为了避免尺度调整过于频繁，或调整时受到噪声干扰。确定 λ_{new} 后，将目标跟踪框放大 λ_{new} 倍作为新的跟踪框。

5 中智加权相似度量尺度自适应跟踪算法流程

将本文提出的跟踪算法称为中智加权相似度

量尺度自适应均值漂移算法，具体流程如下。

步骤 1　在初始帧选取待跟踪目标区域，依据式(8)计算目标特征直方图 $\hat{\boldsymbol{q}}$ 和初始背景直方图 $\boldsymbol{b}$。

步骤 2　读入新的视频帧。

步骤 3　以上一帧的目标位置作为当前帧目标位置的搜寻起始点 $\hat{\boldsymbol{y}}_0$。

步骤 4　以 $\hat{\boldsymbol{y}}_0$ 为起始点，依据式（25）、式（10）和式（19）计算新的目标位置 $\hat{\boldsymbol{y}}_1$。

$$w_i=\sum_{u=1}^{m}\delta\left[b\left(\boldsymbol{x}_i\right)-u\right]\sqrt{w_u^{NSt}\frac{\hat{q}_u}{\hat{p}_u\left(\hat{y}_0\right)}} \tag{25}$$

步骤 5　若不满足 $\left\|\hat{\boldsymbol{y}}_1-\hat{\boldsymbol{y}}_0\right\|<\varepsilon_0$，将 $\hat{\boldsymbol{y}}_0\leftarrow\hat{\boldsymbol{y}}_1$，重复执行步骤 4。

步骤 6　依据式（20）~式（24）计算新的目标尺度，更新目标跟踪框。计算当前帧目标背景特征直方图 $\boldsymbol{b}_c\left(\hat{\boldsymbol{y}}_1\right)$，若 $\rho\left[\boldsymbol{b},\boldsymbol{b}_c\left(\hat{\boldsymbol{y}}\right)\right]<\varepsilon_1$，$\boldsymbol{b}\leftarrow\boldsymbol{b}_c\left(\hat{\boldsymbol{y}}\right)$。

步骤 7　输出目标跟踪框，重复执行步骤 2~步骤 7，直至跟踪结束。

6　实验结果与分析

将本文提出的算法在一些具有挑战性的视频序列中进行了测试，测试序列信息见表 1（其中英文缩写含义复杂背景(BC)、快速移动(FM)、运动模糊（MB）、形变（DEF）、光照变化（IV）、平面内旋转(IPR)、低分辨率(LR)、遮挡(OCC)、平面外旋转（OPR）、尺度变化（SV）），共计 12 396 帧，测试序列中包含了复杂背景、快速移动、运动模糊、形变、光照变化、平面内旋转、低分辨率、遮挡、平面外旋转和尺度变化等挑战。

鉴于提出的算法完全基于均值漂移框架，在选取对比算法时，主要选取了近期基于均值漂移框架表现较优异的算法。为了测试算法性能，将算法与 ASMS[12]、NEUTMS[30]和 KMS[7,8]算法进行了比较。ASMS 是近期表现较好的均值漂移框架下的跟踪算法，它在跟踪的同时能够自适应调整尺度信息；NEUTMS 是基于中智余弦相似度量的均值漂移算法，其在某些方面的性能已超过了 ASMS 算法；鉴于本文算法完全基于 KMS 算法改进，为显示改进效果，也将其引入。除 ASMS 以外，KMS 和 NEUTMS 在跟踪过程中均不更新尺度信息。为方便阐述，下文以 WNEUTSMS 表示本文提出的算法。

表 1　测试视频序列信息

序列名称	跟踪目标	主要挑战	总帧数
David	人脸	IPR, OPR, BC	537
Singer	人	IV, SV, OCC, OPR	351
Bolt	人	DEF, BC	293
Walking	人	SV, OCC, DEF	412
Dog	狗	SV, DEF, OPR	127
Car24	车	IV, SV, BC	3 059
Car1	车	IV, SV, MB, FM, BC, LR	1 020
Car2	车	IV, SV, MB, FM, BC	913
Jump	人	SV, OCC, DEF, MB, FM, IPR, OPR	122
Coke	饮料瓶	IV, OCC, FM, IPR, OPR, BC	291
Lemming	玩偶	IV, SV, OCC, FM, OPR	1 336
Faceocc1	人脸	OCC	892
Faceocc2	人脸	IV, OCC, IPR, OPR	812
Jumping	人	MB, FM	313
Redteam	车	SV, OCC, IPR, OPR, LR	1 918

6.1　参数设置

在测试过程中，所设定的参数值均保持不变。计算式（12）~式（14）时涉及的总帧数 N 被设定为 10；计算背景区域时的扩大系数 λ 被设定为 1.6；式（19）中 w_{C_s}、w_{C_b} 均被设定为 0.5，即两类条件属性在权值向量生成时的贡献等分；式（19）、式（23）中的 α、β、γ 分别被设定为 0.5、0.25、0.25，这里适当强化了真分量的贡献；涉及尺度自适应更新过程缩小和放大系数的 λ_{bsa} 和 λ_{ba} 均被设定为 0.3，式（24）中的比例因子 s 被设定为 1.1；式（24）中的 λ_0 须被设定为一个较小的值，以避免目标框剧烈变化，此处 λ_0=0.04；算法流程中的 ε_0 和 ε_1 分别被设定为 0.1、0.5。

6.2　评价标准

本文主要采用了两种评价方式，一种是中心点误差，该误差是指由算法获得的目标框中心坐标与人为标记的实际目标中心坐标之间的欧氏距离；另

一种是跟踪成功率，其测量基准是跟踪目标区域和实际目标区域的重叠度，该重叠度可表示为：

$$s_i = \frac{\text{area}\left(\text{ROI}_{T_i} \cap \text{ROI}_{G_i}\right)}{\text{area}\left(\text{ROI}_{T_i} \cup \text{ROI}_{G_i}\right)} \tag{26}$$

其中，ROI_{T_i} 为第 i 帧中由跟踪算法获得的目标框区域，ROI_{G_i} 为第 i 帧真实目标框区域。通过设定一个最小重叠度 r，相应的跟踪成功率可表示为：

$$R = \sum_{i=1}^{N} u_i \Big/ N,\ u_i = \begin{cases} 1, s_i > r \\ 0, \text{其他} \end{cases} \tag{27}$$

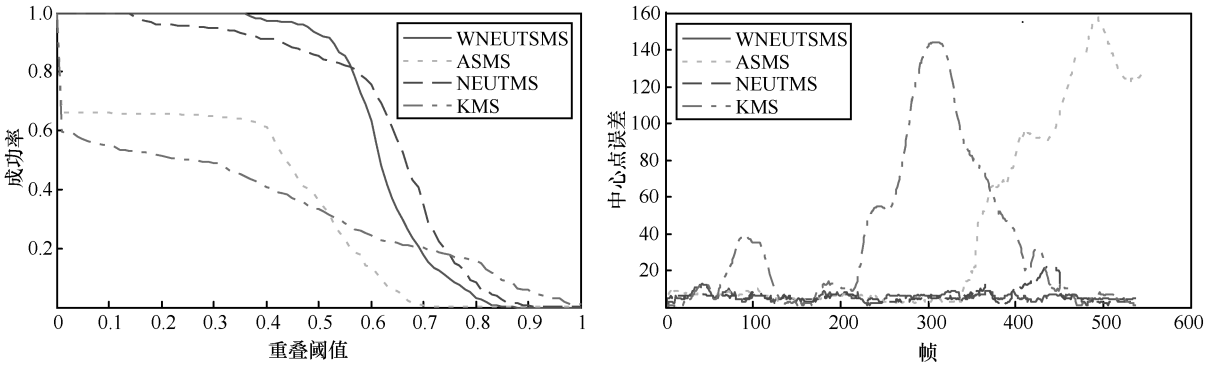
其中，N 为视频序列的总帧数。

6.3 目标跟踪结果

图 2、图 4、图 6、图 8 给出了相应视频序列的跟踪截图，图 1、图 3、图 5、图 7 给出了对应视频序列跟踪结果的成功率曲线和中心点误差曲线，定量分析了几种算法的跟踪效果。接下来针对各序列展开详细分析。

（1）David 序列

该序列中各跟踪算法面临的主要挑战包括目标平面内旋转、平面外旋转和相似背景等。从图 2 可以看到，除了 KMS 算法，其他 3 种跟踪算法在前 330 帧左右表现均较好。在第 65 帧，受到相似



（a）成功率曲线　　（b）中心点误差曲线

图 1　David 序列跟踪性能曲线

图 2　David 序列跟踪结果

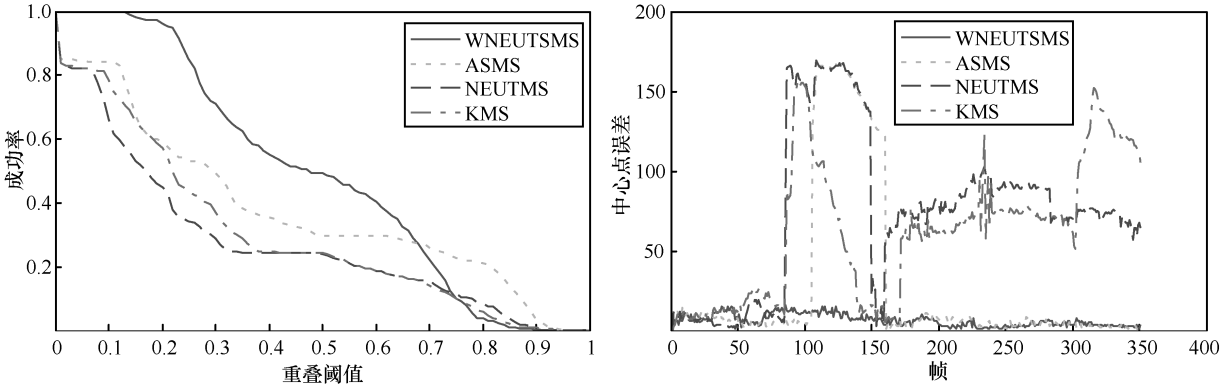
背景区域的干扰，KMS 算法发生偏移。从第 84 帧可以看到 KMS 算法已失败。在第 339 帧，可以看到 ASMS 算法估计了一个过大的尺度，导致目标框较大且开始发生偏移，最终导致跟踪失败（第 360 帧）。在第 430 帧中，受背景干扰和目标旋转影响，NEUTMS 算法发生了较大偏移。综合图 1，可以发现，在此序列中，WNEUTSMS 算法表现最好。

（2）Signer 序列

该序列中包含的挑战包括光照变化、尺度变化等。从图 4 中可以发现，在第 87 帧，除了 WNEUTSMS 和 ASMS 算法以外，其他算法均已远远偏离了目标位置。从尺度估计上，ASMS 在前 87 帧的表现较好，WNEUTSMS 算法给出的目标框相对较小，其原因主要是光照变化导致目标背景发生了剧烈变化，而歌手主体部分仍保持白色主色调，受帧间稳定度约束，目标框锁定在特征较稳定的区域。从第 87 帧开始，仅 WNEUTSMS 算法估计的目标位置准确，且在目标特征恢复时能给出恰当的目标尺度（见第 274 帧、第 322 帧）。观察图 3（b），可以发现虽然 ASMS 算法的中心位置估计尚可，但其估计的目标尺度远大于目标真实尺寸。综合图 3、图 4，可以发现，在本序列中，WNEUTSMS 算法综合表现最优。

（3）Bolt 序列

该序列包含的挑战主要有目标形变、复杂背



（a）成功率曲线　　（b）中心点误差曲线

图 3　Singer 序列跟踪性能曲线

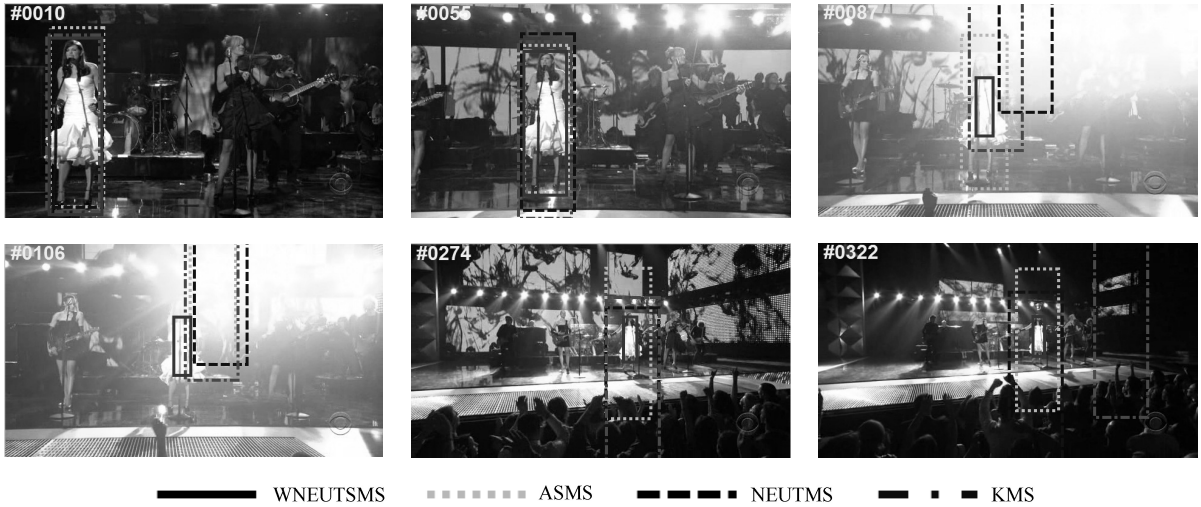
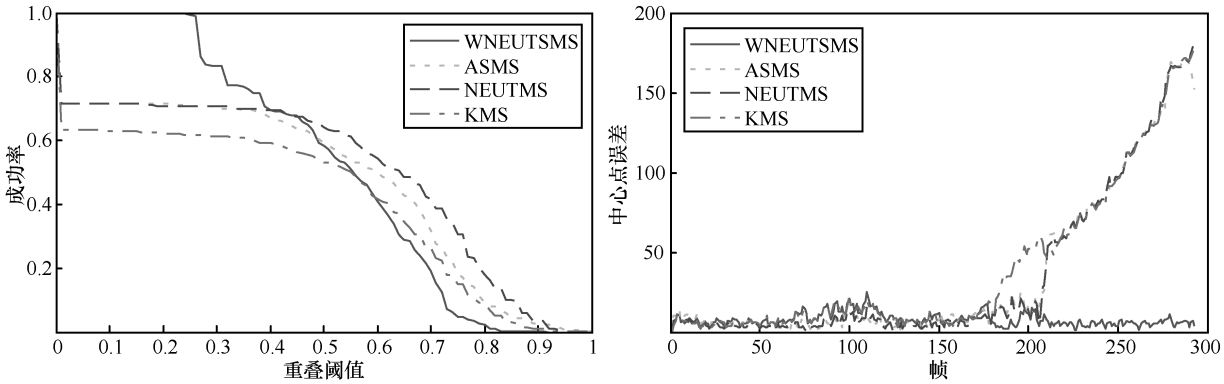


图 4　Singer 序列跟踪结果

景等问题。观察图 5（b），可以发现，从第 200 帧左右开始，除了 WNEUTSMS 算法，ASMS、NEUTMS 和 KMS 算法的跟踪结果均已发生较大偏移，跟踪失败。观察图 6，可以看到在该序列的前半段，4 种算法均能较好的完成目标跟踪。在第 180 帧，算法效果开始出现分化，受 Bolt 旁边运动员外观信息干扰，除 WNEUTSMS 算法外的其他算法受到干扰。借助于尺度估计、帧间稳定特征信息和目标个体特征强化，WNEUTSMS 算法较好地完成了目标跟踪。然而，仔细观察后，会发现 WNEUTSMS 算法估计得到的目标尺度略小于目标成像尺寸，这也是导致在重复率门限较高时，该算法的成功率反而低于其他算法的主要原因。深究造成这一现象的原因，主要是因为目标奔跑速度快、步幅大，导致身体下半部分特征信息极不稳定，故而导致尺度估计偏小。

（4）Walking 序列

该序列中主要包含尺度变化、遮挡、目标形变等挑战。观察图 7（b），可以发现，相比其他 3 种算法，WNEUTSMS 算法在第 150~170 帧、在第 200~230 帧，在第 250~300 帧，均得到了较准确的目标中心点位置。从图 8 中可以看到，在行人接近车辆之前，各算法表现均较好（见第 10 帧、第 143 帧）。在第 154 帧，因行人本身有不少部分呈现黑色特征，在行人靠近黑色车辆时，4 种算



(a) 成功率曲线　　(b) 中心点误差曲线

图 5　Bolt 序列跟踪性能曲线

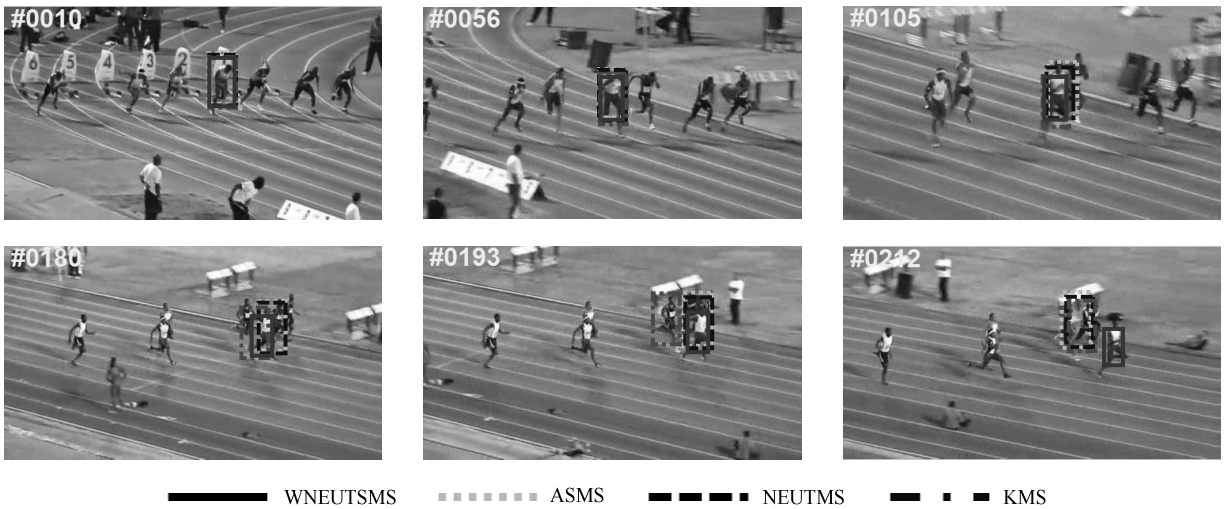
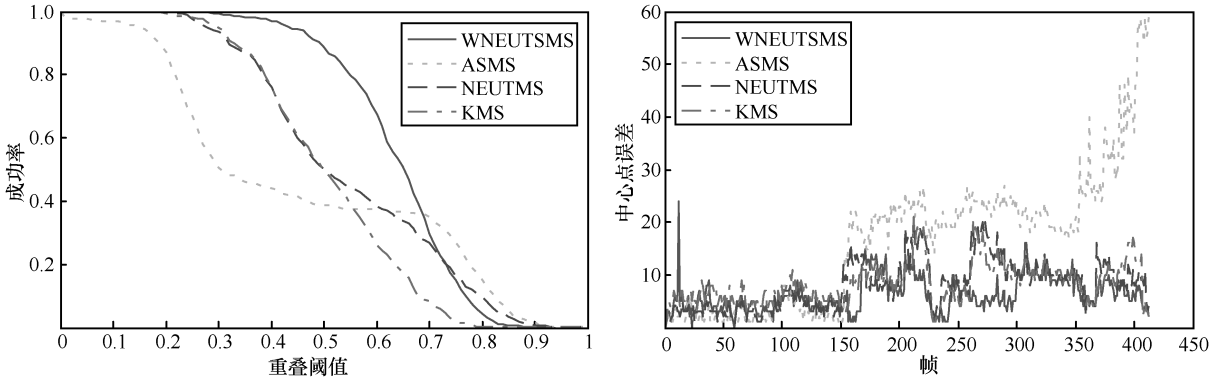


图 6　Bolt 序列跟踪结果

法均不同程度上受到黑色车辆的影响，其中 ASMS 偏离目标中心点最远，WNEUTSMS 算法偏离最小，且估计的目标尺寸最接近目标成像大小。从第 153 帧起，受背景车辆的干扰，ASMS 算法已无法完成尺度信息的准确估计，故而逐步导致算法失败，而 WNEUTSMS 算法能够给出相对合理的目标框尺寸，这在图 7（a）成功率曲线上也有所体现。

为评价本文算法在所有测试序列中的表现，将 TRE 和 OPE 评价引入。为完成 TRE 评价，每个测试序列被分为 20 份序列片段（各序列片段对应原序列不同起始帧，但均以原序列结束帧为终止），每一个跟踪算法均在每份测试序列片段中进行测试。对于 OPE 评价，每个测试序列是从第一帧测试到最后一帧，且仅测一次。图 9 和图 10 分别给出了基于 TRE、OPE 评价的面向表 1 所示整个测试序列集的测试结果，表 2 和表 3 分别给出了各跟踪算法应对不同挑战的 TRE-AUC、OPE-AUC 分值。AUC 分值为成功率曲线中对应成功率曲线与横纵坐标轴围成区域的面积，该值越大，表明算法越顽健。TRE-AUC 分值与 TRE 成功率曲线对应，OPE-AUC 分值则与 OPE 成功率曲线对应。



（a）成功率曲线　　（b）中心点误差曲线

图 7　Walking 序列跟踪性能曲线

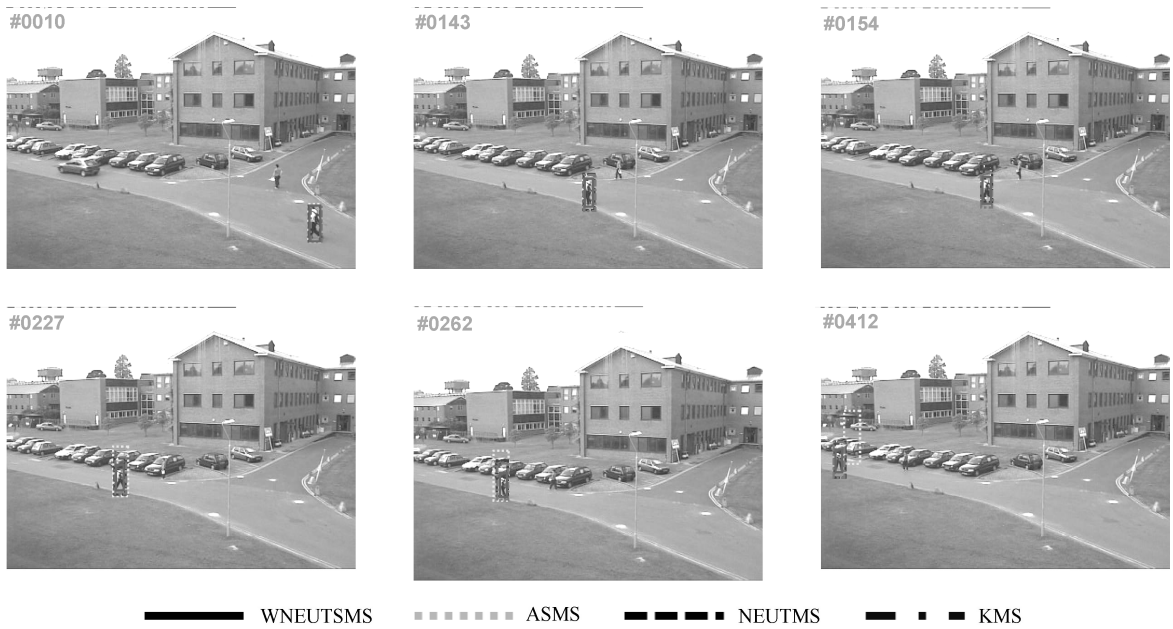


图 8　Walking 序列跟踪结果

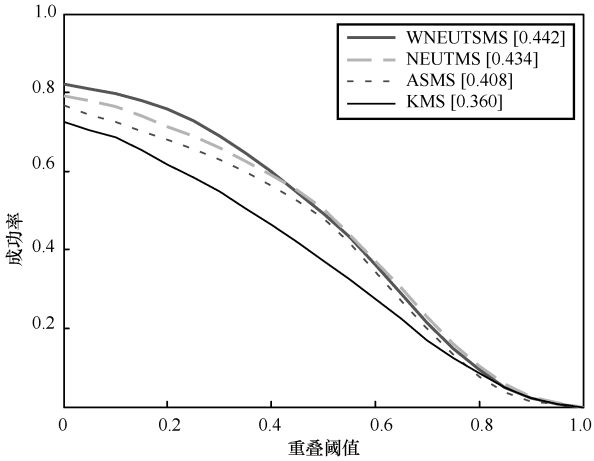


图 9　所有测试序列 TRE 成功率曲线

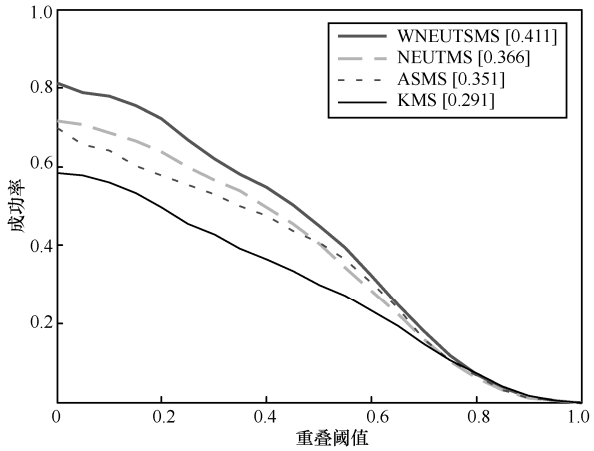


图 10　所有测试序列 OPE 成功率曲线

综合分析图 9、图 10 和表 2、表 3，WNEUTSMS 算法总体表现较好。观察表 2，WNEUTSMS 算法在应对复杂背景、光照变化、平面内旋转、遮挡和平面外旋转时表现最好；在应对形变挑战时仅次于 NEUTMS 算法，其原因主要在于大范围的形变可能会造成目标跟踪框尺度估计不准确；在应对低分辨率和尺度变化时的表现仅次于 ASMS 算法，其原因主要在于在低分辨率情况下，仅有较少的信息能够被用来提升算法性能。对于 OPE 评价，WNEUTSMS 算法在应对快速移动、运动模糊、形变、平面内旋转、遮挡和平面外旋转时表现最好。在应对复杂背景、光照变化时仅次于 ASMS 算法。

综合以上分析，可以发现，本文算法（WNEUTSMS 算法）在确定目标位置时将目标帧间特征稳定性和背景信息引入，同时强化“真”量测分量；在确定目标尺度时联合考虑目标特征信息和目标背景信息。多类信息融合考量提升了算法在应对背景干扰、光照变化、尺度变化等挑战时的顽健性。

7　结束语

本文提出了一种面向单值中智集的分量加权

表 2　各跟踪算法应对不同挑战的 TRE-AUC 分值

挑战	BC	FM	MB	DEF	IV	IPR	LR	OCC	OPR	SV	综合
WNEUTSMS	0.303	0.477	0.399	**0.450**	0.371	0.445	**0.290**	0.469	0.466	**0.415**	0.442
NEUTMS	**0.279**	**0.485**	0.452	0.459	**0.354**	**0.435**	0.253	**0.461**	**0.460**	0.387	**0.434**
ASMS	0.253	0.513	**0.407**	0.403	0.338	0.417	0.349	0.455	0.453	0.445	0.408
KMS	0.242	0.366	0.265	0.366	0.329	0.342	0.199	0.438	0.382	0.337	0.360

注：表 2 中下划线数值为对应列最大值，加粗数值为对应列次最大值。

表 3　各跟踪算法应对不同挑战的 OPE-AUC 分值

挑战	BC	FM	MB	DEF	IV	IPR	LR	OCC	OPR	SV	综合
WNEUTSMS	**0.253**	0.391	0.377	0.470	**0.319**	0.372	0.250	0.450	0.424	0.426	0.411
NEUTMS	0.187	**0.390**	**0.325**	**0.363**	0.248	**0.349**	**0.279**	0.379	**0.374**	0.311	**0.366**
ASMS	0.282	0.364	0.126	0.327	0.321	0.297	0.377	0.375	0.358	**0.397**	0.351
KMS	0.161	0.218	0.081	0.325	0.265	0.238	0.087	**0.397**	0.298	0.257	0.291

注：表 3 中加下划线数值为对应列最大值，加粗数值为对应列次最大值。

余弦相似度量，并基于此在中智框架内提升了均值漂移视觉跟踪算法的顽健性。针对目标帧间特征稳定性和目标/背景相似度两种属性，本文提出了相应的真、不确定、假量测计算方式，然后利用分量加权余弦相似度量构建了新的权值向量，用于修正传统均值漂移跟踪算法的权值。此外，本文首次提出了基于中智加权余弦相似度量的尺度更新算法。实验结果表明，本文提出的视觉跟踪算法能较好应对相似背景、光照变化和尺度变化等挑战。

参考文献：

[1] 赵利平, 林涛, 周开伦, 等. 一种多色度格式级联编码的 AVS2 全色度图像编码算法[J]. 电信科学, 2018, 34(4): 57-67.
ZHAO L P, LIN T, ZHOU K L, et al. A multi-chroma format cascaded coding method for full-chroma image in AVS2[J]. Telecommunications Science, 2018, 34(4): 57-67.

[2] 唐彪, 金炜, 符冉迪, 等. 多稀疏表示分类器决策融合的人脸识别[J]. 电信科学, 2018, 34(4): 31-40.
FU B, JIN W, FU R D, et al. Face recognition using decision fusion of multiple sparse representation-based classifiers[J]. Telecommunications Science, 2018, 34(4): 31-40.

[3] 颜文, 金炜, 符冉迪. 结合 VLAD 特征和稀疏表示的图像检索[J]. 电信科学, 2016, 32(12): 80-85.
YAN W, JIN W, FU R D. Image retrieval based on the feature of VLAD and sparse representation[J]. Telecommunications Science, 2016, 32 (12): 80-85.

[4] YILMAZ A, JAVED O, SHAH M. Object tracking: a survey [J]. ACM Computing Surveys, 2006, 38(4): 13.

[5] WU Y, LIM J, YANG M H. Online object tracking: a benchmark[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23-28, 2013, Portland, OR, USA. Piscataway: IEEE Press, 2013: 2411-2418.

[6] SMEULDERS A W M, CHU D M, CUCCHIARA R, et al. Visual tracking: an experimental survey [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(7): 1442-1468.

[7] COMANICIU D, RAMESH V, MEER P. Real-time tracking of non-rigid objects using mean shift[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 13-15, 2000, Hilton Head Island, SC, USA. Piscataway: IEEE Press, 2000: 142-149.

[8] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.

[9] LEICHTER I. Mean shift trackers with cross-bin metrics[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(4): 695-706.

[10] ZHOU H, YUAN Y, SHI C. Object tracking using SIFT features and mean shift[J]. Computer Vision and Image Understanding, 2009, 113(3): 345-352.

[11] BOUSETOUANE F, DIB L, SNOUSSI H. Improved mean shift integrating texture and color features for robust real time object tracking[J]. The Visual Computer, 2013, 29(3): 155-170.

[12] VOJIR T, NOSKOVA J, MATAS J. Robust scale-adaptive mean-shift for tracking[J]. Pattern Recognition Letters, 2014(49): 250-258.

[13] COLLINS R T. Mean-shift blob tracking through scale space[C]//IEEE Conference on Computer Vision Pattern Recognition (CVPR), June 16-22, 2003, Madison, WI, USA. Piscataway: IEEE Press, 2003: 234-234.

[14] SMARANDACHE F. Neutrosophy: neutrosophic probability, set and logic[M]. Rehoboth: American Research Press, 1998: 105.

[15] WANG H, SMARANDACHE F, ZHANG Y Q, et al. Single valued neutrosophic sets[J]. Multispace and Multistructure, 2010(4): 410-413.

[16] YE J. Vector similarity measures of simplified neutrosophic sets and their application in multicriteria decision making[J]. International Journal of Fuzzy Systems, 2014, 16(2): 204-211.

[17] GUO Y, ŞENGÜR A. A novel image segmentation algorithm based on neutrosophic similarity clustering[J]. Applied Soft Computing Journal, 2014(25): 391-398.

[18] GUO Y, XIA R, ŞENGÜR A, et al. A novel image segmentation approach based on neutrosophic C-means clustering and indeterminacy filtering[J]. Neural Computing and Applications, 2016: 1-11.

[19] HU K, YE J, FAN E, et al. A novel object tracking algorithm by fusing color and depth information based on single valued neutrosophic cross-entropy[J]. Journal of Intelligent and Fuzzy Systems, 2017, 32(3): 1775-1786.

[20] GUO Y, SENGUR A. A novel 3D skeleton algorithm based on neutrosophic cost function[J]. Applied Soft Computing Journal, 2015(36): 210-217.

[21] 崔西希, 吴成茂. 核空间中智模糊聚类及图像分割应用[J]. 中国图象图形学报, 2016, 21(10): 1316-1327.
CUI X X, WU C M. Neutrosophic C-means clustering in kernel space and its application in image segmentation[J]. Journal of Image and Graphics, 2016, 21(10): 1316-1327.

[22] 张桂梅, 王大雷. 结合 LPG&PCA 的中智学图像分割[J]. 中国图象图形学报, 2014, 19(5): 693-700.
ZHANG G M, WANG D L. Neutrosophic image segmentation approach integrated LPG&PCA [J]. Journal of Image and Graphics, 2014, 19(5): 693-700.

[23] 郑肇葆, 潘励, 郑宏. 中智逻辑图像分割方法的研究与分析[J].

武汉大学学报(信息科学版), 2015, 40(2): 143-146.
ZHENG Z B, PAN L, ZHENG H. Research and analysis of neutrosophic logic image segmentation (NLIS) method[J]. Geomatics and Information Science of Wuhan University, 2015, 40(2): 143-146.

[24] YE J. Single-valued neutrosophic similarity measures based on cotangent function and their application in the fault diagnosis of steam turbine[J]. Soft Computing, 2015, 21(3): 817-825.

[25] MA Y X, WANG J Q, WANG J, et al. An interval neutrosophic linguistic multi-criteria group decision-making method and its application in selecting medical treatment options[J]. Neural Computing and Applications, 2016: 1-21.

[26] YE J, FU J. Multi-period medical diagnosis method using a single valued neutrosophic similarity measure based on tangent function[J]. Computer Methods & Programs in Biomedicine, 2016(123): 142-149.

[27] ANTER A M, HASSANIEN A E, ELSOUD M A A, et al. Neutrosophic sets and fuzzy C-means clustering for improving CT liver image segmentation[J]. Advances in Intelligent Systems and Computing, 2014(303): 193-203.

[28] GUO Y, SENGUR A. NECM: neutrosophic evidential C-means clustering algorithm[J]. Neural Computing and Applications, 2015, 26(3): 561-571.

[29] GUO Y, SENGUR A. NCM: Neutrosophic c-means clustering algorithm[J]. Pattern Recognition, 2015, 48(8): 2710-2724.

[30] HU K, FAN E, YE J, et al. Neutrosophic similarity score based weighted histogram for robust mean-shift tracking[J]. Information, 2017, 8(4): 122.

[31] YE J. Multicriteria decision-making method using the correlation coefficient under single-valued neutrosophic environment[J]. International Journal of General Systems, 2013, 42(4): 386-394.

[作者简介]

胡珂立（1989−），男，博士，绍兴文理学院计算机科学与工程系讲师，主要研究方向为图像处理、目标跟踪、中智理论。

范恩（1982−），男，博士，绍兴文理学院计算机科学与工程系讲师，主要研究方向为传感器数据融合、目标跟踪、中智理论。

叶军（1959−），男，绍兴文理学院电子与信息工程系教授，主要研究方向为中智理论。

沈士根（1974−），男，博士，绍兴文理学院计算机科学与工程系教授，主要研究方向为无线传感器网络、移动互联网、博弈论。

谷宇章（1976−），男，博士，中科院上海微系统与信息技术研究所物联网系统技术实验室副研究员，主要研究方向为图像处理、仿生机器人视觉、立体视觉感知。