

基于增强学习的摄像机网络节点动态选择方法*

李 骞^{1,2}, 孙正兴¹, 陈松乐¹, 夏士明²

¹(计算机软件新技术国家重点实验室(南京大学), 江苏 南京 210093)

²(解放军理工大学 气象海洋学院, 江苏 南京 211101)

通讯作者: 孙正兴, E-mail: szx@nju.edu.cn, http://cs.nju.edu.cn/szx/

摘要: 摄像机节点动态选择问题是摄像机网络应用中的一个难点. 提出了一种基于增强学习的节点动态选择方法. 采用视觉信息评分作为单步回报设计了节点选择策略的 Q -学习算法, 为了加速算法收敛速度, 利用摄像机空间拓扑关系初始化 Q 值表, 并基于 Gibbs 分布进行非贪心尝试. 从目标可见性、朝向、清晰度和切换次数设计视觉评价函数反映视频信息丰富程度和视觉舒适度. 实验结果表明, 该节点动态选择方法能够有效地反映视频中的目标状态信息, 选择结果切换平滑, 满足实际应用需要.

关键词: 摄像机选择; 增强学习; 视频分析; 摄像机网络

中文引用格式: 李骞, 孙正兴, 陈松乐, 夏士明. 基于增强学习的摄像机网络节点动态选择方法. 软件学报, 2015, 26(Suppl. (2)): 8-19. <http://www.jos.org.cn/1000-9825/15011.htm>

英文引用格式: Li Q, Sun ZX, Chen SL, Xia SM. Node dynamic selection in camera networks based on reinforcement learning. Ruan Jian Xue Bao/Journal of Software, 2015, 26(Suppl. (2)): 8-19 (in Chinese). <http://www.jos.org.cn/1000-9825/15011.htm>

Node Dynamic Selection in Camera Networks Based on Reinforcement Learning

LI Qian^{1,2}, SUN Zheng-Xing¹, CHEN Song-Le¹, XIA Shi-Ming²

¹(State Key Laboratory for Novel Software Technology (Nanjing University), Nanjing 210093, China)

²(College of Meteorology and Oceanography, PLA University of Science and Technology, Nanjing 211101, China)

Abstract: This paper addresses the problem of node dynamic selection in camera networks. A selection method based on reinforcement learning is proposed in which the node is selected to maximize the expected reward while minimizing the switching with Q -learning. To accelerate the convergence of Q -learning, the geometry of camera networks is considered for initial Q -values and a Gibbs distribution is used for exploitation. In order to evaluate visual information of the video, a function of the visibility, orientation, definition and switching is designed to assess the immediate reward in Q -learning. Experiments show that the proposed visual evaluation criteria can capture the motion state of the object effectively and the selection method is more accurate on reducing cameras switching compared with the state-of-the-art methods.

Key words: camera selection; reinforcement learning; video analysis; camera network

近年来, 摄像机网络成为计算机视觉及应用领域的热点之一. 在有效解决单摄像机存在的视域(field of view)范围有限、遮挡和持续跟踪等问题的同时, 摄像机网络也带来了诸如多摄像机部署和控制、多路高分辨率和高帧率视频流实时融合、多摄像机协作和节点选择等新的挑战^[1-3], 其中, 摄像机节点的自动选择, 即从多路视频中选取并推送富信息量的视频, 成为基于摄像机网络的目标跟踪与定位^[4-6]、区域覆盖^[7]、监控与行为分析^[8,9]和视频摘要^[10]等应用的基本问题之一, 尤其在近年来不断涌现的智能家庭、关键区域安全监控等应用中,

* 基金项目: 国家自然科学基金(61272219, 61100110, 61321491, 41305138); 教育部新世纪优秀人才资助计划(NCET-04-0460); 江苏省科技计划(BE2010072, BE2011058, BY2012190, BY2013072-04); 计算机软件新技术国家重点实验室创新基金(ZZKT2013A12)

收稿时间: 2014-01-03; 定稿时间: 2014-04-18

通过节点选择获取特定目标高质量视频已成为解决视频数据传输、节点供电和显示终端数目等条件受限问题的有效途径。与一般传感器网络节点选择问题^[11]不同的是,摄像机感知环境具有方向性,不能简单地根据目标的位置远近选择最优节点,同时受到背景、光照等因素影响,视觉计算引入了一定的误差,处理结果在时间序列上表现出一定的不稳定性。与之类似,针对三维模型的视角选择问题^[12,13]在图形学领域中也进行了广泛研究,与摄像机节点选择问题不同的是,前者要求精确定义对象模型,无需考虑背景等因素,且模型分析通常在静态环境下进行。综上所述,摄像机网络中节点选择问题难点在于:首先要建立动态选择机制,即应在某一时间点上基于当前和历史观察选择合适节点,使期望视觉信息量最大,且输出视频应尽量平滑无频繁切换;其次需要根据应用需求设计可适应不同摄像机参数、背景和环境变化的视觉信息评价准则。

现有的摄像机网络节点选择方法主要分为 3 类。第 1 类方法仅以当前视觉信息量最大为目标进行贪心选择,如文献[8,14–15]等。这类方法既不考虑选择后视觉信息的状态变化,也不考虑切换、通信、能耗等代价。第 2 类方法设计一定效用或者代价函数以融合多个约束条件,对选择过程进行优化,如文献[16–18]等。这类方法仅针对当前一定时间窗内状态进行优化,不考虑选择回报延迟。第 3 类方法则采用调度的思想,基于过去或当前状态做出动态决策,如文献[19]利用动态贝叶斯网实现视频产品的离线最优视角选择,文献[20–22]分别采用部分可观察马尔可夫模型(POMDP)对跟踪与监控应用中摄像机调度策略进行建模和求解。这类方法预测了选择动作对未来视觉信息评价产生的影响,但上述方法均需明确状态转移模型,而动态场景中系统状态变化难以显式地定义,计算复杂度也将随状态空间呈指数级增长,易引起“维度灾难”^[23],制约了在复杂监控场景下的应用。

另一方面,对视频信息量的评价主要通过提取有效视觉特征并将其映射为可比较的评分来实现。其中合适视觉特征的选择是评价的关键,与应用需求紧密相关,如跟踪与定位应用中常采用目标可见性^[24]、图像或真实世界位置^[14–15,20]、目标图像区域大小^[16]等视觉特征,监控与行为分析应用更关注监控对象的面部朝向^[8]、目标姿态与外观^[9,25]和场景运动的数量^[19]。针对关键区域的监控应用一般要求获取感兴趣目标正面且清晰度较高的视频,同时应尽可能地避免因视觉计算误差产生的频繁切换。尽管通过姿态估计和人脸检测可获得信息量较大的正面图像,但从单一视角估计人体姿态是一个病态和非线性问题^[26],人脸检测成功率也容易受视频分辨率及目标摄像机相对位置的影响,两种方法计算复杂度高且输出结果不稳定。

基于以上分析,本文提出了一种面向关键区域监控应用的摄像机动态选择方法,在状态变化未知的动态场景中,融合视域内特定目标的可见性、运动方向和视频画面质量等特征,采用增强学习相关方法在线学习最优节点选择策略,以达到视觉信息量最大化与视角切换最少等约束条件的平衡。

1 基于 Q -学习的选择方法

1.1 问题定义

本文所述摄像机网络节点选择问题如图 1 所示,特定场景下部署 N 个静止摄像机 (C^1, C^2, \dots, C^N), 摄像机节点之间部分或全部重叠视域,存在一个中心节点负责摄像机的选择与调度。目标在场景中运动时,在选择时间点 $t = 0, 1, 2, \dots$, 中心节点从已选摄像机获取的视频对目标运动、朝向及视频质量进行分析,选择最优摄像机 C^* , 在时间段 $(t, t + \Delta t)$ 内由 C^* 输出目标监控视频,以获取选择动作的最大期望回报。

在上述选择过程中,为了获取高质量的目标监控视频,中心节点每次需根据已有视频和视觉计算得到的目标状态动态选择最优摄像机,因此可视为一个基于视觉内容分析和评价的序列决策过程。由于选择过程中状态转移和期望回报仅取决于当前状态和选择动作,可建模为一个马尔可夫决策过程(Markov decision processes, 简称 MDP),采用四元组 $\langle A, S, T, R \rangle$ 表示, A 为可选摄像机集合 $\{C^1, C^2, \dots, C^N\}$; S 定义状态空间,对第 t 个选择决策点上的系统状态 $s_t \in S$, 本文表示为向量 $s_t = (P_t, O_t, a_{t-1})$, 其中 P_t 为上一选择点 $t-1$ 至当前选择点 t 共 K 个时刻的目标的位置序列 $\{(x_t^1, y_t^1), (x_t^2, y_t^2), \dots, (x_t^K, y_t^K)\}$, O_t 为目标朝向序列 $\{\theta_t^1, \theta_t^2, \dots, \theta_t^K\}$, $a_{t-1} \in A$ 为 $t-1$ 选择点所选节点, T 为状态转移函数 $T: S \times A \rightarrow S$, $R(s_t, a_t)$ 为在第 t 个选择点的单步回报函数 $R: S \times A \rightarrow \mathbb{R}$, 通过计算最优策略 $\pi^*(s): s \mapsto a$, 使该状态下节点选择的期望总回报值函数 $V^{\pi^*}(s)$ 最大,即

$$V^{\pi^*}(s) = \max_{\pi} E \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \mid s_0 = s \right] \quad (1)$$

其中, γ 为折扣系数 $0 \leq \gamma < 1$. 当状态转移模型和回报函数已知时, 上述模型可采用动态规划离线求解. 但在复杂动态环境下, 转换关系难以显式地定义, 因此本文提出了基于 Q -学习方法, 对选择问题进行建模与求解.

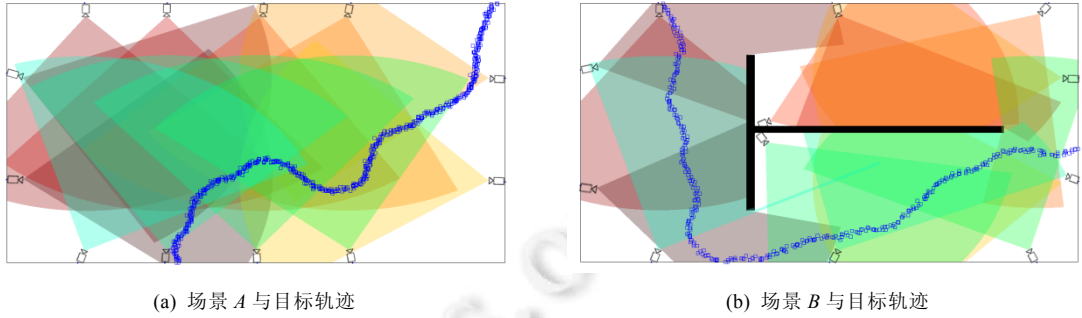


Fig1. Example of scene and target's trajectory

图 1 场景和目标轨迹示例

1.2 算法框架

Q -学习是一种重要的无模型增强学习方法^[27], 可在未知动态环境中通过估计优化动作值函数求取最优策略. 本文提出了一种基于 Q -学习的摄像机网络节点选择方法, 学习数据由多个插曲(episodic)形式数据序列组成, 学习一个插曲结束后开始新的插曲学习. 每个插曲以目标离开监控区域或选择动作次数达到指定的数目 T 为结束标志, 插曲数据由状态、选择动作和回报序列表示 $s_0, a_0, r_1, s_1, \dots, s_t, a_t, r_{t+1}, s_{t+1}, \dots, s_{T-1}, a_{T-1}, r_T, s_T$. 其中, s_t 为插曲中第 t 个选择点的系统状态, a_t 为在已有策略指导下基于当前状态的所选摄像机节点, r_{t+1} 为摄像机选择后一段时间内所获取的单步选择回报. 单个插曲训练后, 对插曲内每组状态-动作对 (s_t, a_t) 增量估计对应 Q 函数值 $Q(s_t, a_t)$, 基于 Q -学习时间差分思想可定义更新准则如下:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (2)$$

其中, α 为学习率, 反映新值对 Q 函数值更新速率影响程度, 取值范围为 $[0, 1]$; γ 为期望回报计算时的折扣因子; $\max_a Q(s_{t+1}, a)$ 为对观察到的后继状态选择动作 a 使 Q 函数值最大. 文献^[27]证明, 当 α 满足一定条件时 Q -学习算法将收敛于最优状态-动作函数. 待 Q 函数趋向收敛时, 可对学习后的 Q 函数通过最大化求取最优策略:

$$\pi(s) = \arg \max_{a \in A} Q(s, a) \quad (3)$$

因此, 本文算法可由算法 1 来表示.

算法 1. Camera selection based Q -learning.

Inputs:

γ : Discount factor, α : initial learning parameter

S : states space, A : camera sets

Outputs:

π : The Optimal policy for camera selection

BEGIN:

1. Initialize $Q(s, a) = Q_0(s, a)$ for $\forall s \in S, \forall a \in A$

2. **FOR each** episode **DO**

3. initialized s_t as starting state s_0

4. **REPEAT**
5. selection the camera $a_t \in A$ as follow rule
6.
$$a_t \leftarrow \begin{cases} a := \arg \max_a (Q(s_t, a)), & \text{with the probability } 1 - \varepsilon \\ \text{a Gibbs random camera selection,} & \text{with the probability } \varepsilon \end{cases}$$
7. turn on camera a_t and turn off the other cameras
8. observe the new state s_{t+1} with the camera a_t and get the immediate reward $r_{t+1}(s_{t+1}, a_t, a_{t+1})$
9. update $Q(s, a)$ as $Q(s_t, a_t) := Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$
10. $s_t := s_{t+1}, a_t := a_{t+1}, t := t + 1$
11. **UNTIL** episode is end
12. **END**
13. Find the optimal policy π as $\pi(s) = \arg \max_{a \in A} Q(s, a)$

END

其中,即时回报 r_{t+1} 为选择节点 a_t 后至下一选择结果 a_{t+1} 的时段内目标视觉信息评分,将在第 2 节加以介绍.

1.3 Q 值初始化

大状态空间的学习收敛速度一直是 Q -学习算法需要研究解决的问题.在本文中若目标后续状态保持稳定不变,选择依据应为当前视觉信息量最大,因此为加速学习过程的收敛,算法 1 利用目标状态与摄像机拓扑作为先验知识,将初始 Q 值设为贪心选择摄像机的即时回报 $Q_0(s, a)$, 本文假设在摄像机硬件配置相同条件下根据目标与摄像机的位置与朝向关系对 $Q_0(s, a)$ 定义,即

$$Q_0(s, a) = w_0 \left(\sin \left(\frac{|\theta_s - \theta_a|}{2} \right) \right) + (1 - w_0) \left(1 - \frac{\sqrt{(x_s - x_a)^2 + (y_s - y_a)^2}}{\text{MaxDis}} \right) \quad (4)$$

其中, w_0 为权值, (x_s, y_s) 与 (x_a, y_a) 分别为目标和摄像机 a 位置, MaxDis 为归一化距离, θ_a 和 θ_s 分别表示所选摄像机 a 和目标在状态 s 下正面的朝向角.上式表明,当目标距离摄像机越近,目标与摄像机夹角越小时,可得到质量更好的监控画面.

1.4 动作尝试策略

增强学习本质是通过 Agent 不断与外部环境的交互获取变化模型等知识,而交互包括演算(exploitation)和尝试(exploration),分别是在状态一定条件下按已学习的策略选择动作和尝试新的动作对问题空间进行探索,其中尝试策略是获取对问题空间新知识的关键,一般采用 ε -greedy 策略^[28].本文算法 1 中第 6 行仍采用 ε -greedy 策略,以概率 $1 - \varepsilon$ 基于当前 Q 值贪心选择摄像机,即 $a_t = \arg \max_a Q(s_t, a)$.为使较大 Q 值的候选摄像机在非贪心尝试时具有较大的被选概率,将 ε -greedy 策略中的随机选择改进为基于 Gibbs 分布的节点选择,在 t 时刻节点 a_t 被选概率 p_{a_t} 定义为

$$p_{a_t} = \frac{e^{Q(s_t, a_t)\varepsilon}}{\sum_{b=1}^N e^{Q(s_t, b)\varepsilon}} \quad (5)$$

其中, ε 为设定的尝试概率, N 为候选摄像机节点数量.

2 视觉信息评价

第 1 节中单步选择回报由视觉信息评分表示.我们从视频中单帧图像获取的目标信息与输出视频的视觉舒适度要求出发,将选择回报 r_{t+1} 分为选择间隔时段内单帧图像视觉评分均值(正回报)与视角切换代价(负回报)两部分,如式(6)所示.

$$r_{t+1} = w_0 \overline{M_{t+1}^{a_t}} - w_1 \delta(a_t, a_{t+1}) \quad (6)$$

$$\overline{M}_{t+1}^{a_t} = \frac{1}{K} \sum_{k=t+1}^{t+K} M_k^{a_t} \tag{7}$$

$$\delta(a_t, a_{t+1}) = \begin{cases} 0, & a_t = a_{t+1} \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

式(6)中, w_0 和 w_1 为归一化权重, $\overline{M}_{t+1}^{a_t}$ 为在第 t 个选择点选择摄像机 a_t 后 K 幅图像的视觉信息评分均值, 如式(7)所示. 式(8)中 $\delta(a_t, a_{t+1})$ 为相邻选择点的 δ 函数. 某时刻摄像机 a_t 单帧图像目标视觉信息评分 M^{a_t} 可由目标可见性评分 $M_v^{a_t}$, 归一化的朝向评分 $M_d^{a_t}$ 与目标图像清晰度评分 $M_c^{a_t}$ 表示

$$M^{a_t} = M_v^{a_t} (w_2 M_d^{a_t} + w_3 M_c^{a_t}) \tag{9}$$

其中, w_2 和 w_3 分别为对应评分分量的权重, 其大小关系反映实际应用对朝向信息和视频中目标清晰度的关注程度, 不同应用可对权重进行调整. 在考察摄像机视觉特征及评分时, 为简化问题, 我们将目标可见性、朝向和图像清晰度等评分分量以摄像机无关形式表示, 即 M_v, M_d 和 M_c .

为计算上述评分分量, 首先对被选摄像机 a_t 中的目标进行跟踪, 得到目标的外接矩形图像序列, 由外接矩形底部中心坐标与该摄像机标定参数可得到目标在客观世界的坐标序列 P_{t+1} . 因此, 各评分分量计算方法如下.

1) 可见性评分. 监控过程中能否捕获场景中目标成为评价摄像机选择结果的重要依据. 本文主要通过运动检测过程中最大运动区域面积推测目标在特定视角下的可见性, 其主要思想是, 当最大运动区域的面积小于一定阈值时, 可推测该视角下目标不可见, 即 $M_v = 0$, 否则 $M_v = 1$.

2) 朝向评分. 本文利用目标运动信息计算目标朝向信息, 假定目标正面朝向与其运动方向一致且在较小时间范围内目标作近似直线运动, 对跟踪得到的位置序列 P_{t+1}, k 时刻目标朝向角 θ_t^k 可由 $\{(x_t^{k-l+1}, y_t^{k-l+1}), (x_t^{k-l+2}, y_t^{k-l+2}), \dots, (x_t^k, y_t^k)\}$ 共 l 个位置序列直线拟合得到的斜率表示, 如图 2 所示. 通过直线的拟合, 减少了由标定和跟踪算法引入的视觉计算误差. 与式(4)相似, 本文选定视角下 k 时刻目标朝向评分可定义为

$$M_d = \sin\left(\frac{|\theta_t^k - \theta_a|}{2}\right) \tag{10}$$

上式表明, 目标运动方向正对所选摄像机, 即夹角为 180° 时朝向评分最大.

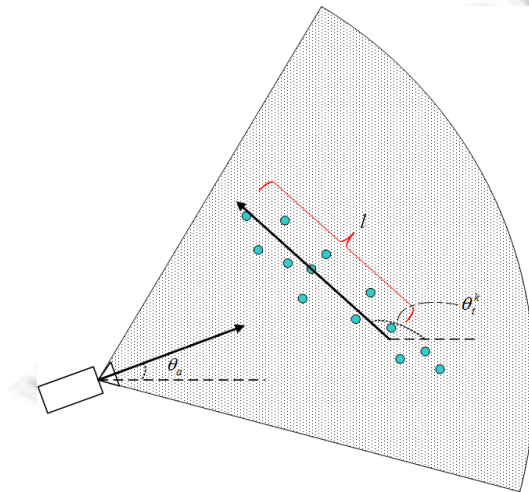


Fig.2 Orientation angle based on target's position sequence

图 2 目标位置序列计算朝向角

3) 清晰度评分. 不同视角下目标图像清晰程度反映了摄像机获取细节信息能力的大小, 通常图像清晰度可由平均梯度或信息熵等形式表示. 由于目标在不同视角下的成像大小不同, 为比较不同视角下目标图像的清晰

度,本文将检测到的目标区域图像按比例放大到高为 H 的图像,若放大后图像宽为 W ,则目标在该视角下的清晰度评分 M_c 可由缩放后图像的平均梯度表示,即

$$M_c = \frac{1}{(W-1)(H-1)} \sum_{x=1}^{W-1} \sum_{y=1}^{H-1} \sqrt{\left(\frac{\partial I(x,y)}{\partial x}\right)^2 + \left(\frac{\partial I(x,y)}{\partial y}\right)^2} \quad (11)$$

其中, $I(x,y)$ 为坐标 (x,y) 处的灰度值.

3 实验结果

3.1 实验建立

本文在配置为 Intel Core i7-3770(CPU),8G 内存的 PC 机上设计并实现了实验.第 3.2 节设计场景和摄像机参数不同条件下带有噪声的仿真实验,对比贪心选择结果验证本文方法的收敛性和有效性,采用仿真实验可有效弥补真实实验中硬件与场地受限等不足.第 3.3 节在真实场景中部署摄像机,对本文视觉信息评价特征进行测试与分析,同时选择结果与基于贪心准则的当前最大视觉评分值法(Max)、博弈论方法(GAME)^[17]、部分可观察马尔可夫模型(POMDP)^[21]进行对比.方法中学习率和尝试概率参照一般增强学习方法分别设为 $\alpha = 0.1, \varepsilon = 0.1$,每个插曲由 $T = 10$ 个选择点组成,设折扣因子 $\gamma = 0.7$,此时 $\gamma^T < 0.02$,可认为当前选择动作不会对 T 以后的状态产生影响.如第 2 节所述,特征权重反映应用需求对视觉特征关注程度,本文设定为 $w_0 = 0.5, w_1 = 0.5, w_2 = 0.6, w_3 = 0.4$.为在保持方向变化增加朝向信息评价中直线拟合的稳定性,设拟合合点数 l 为 10,此外被选节点目标图像按比例放大成高 H 为 300 的图像后计算清晰度评分.

3.2 仿真实验结果与分析

我们将 12 个静止摄像机随机部署在长为 12 000mm,宽为 8 000mm,高为 2 000mm 的虚拟区域,并进行了大量仿真实验,本文以图 1 中 A, B 两个典型场景为例进行仿真实验分析.在不同场景中模拟单个目标各生成 20 000 组运动轨迹作为训练数据,目标状态由位置坐标 (x, y) 和朝向角 (θ) 组成,在真值数据基础上对每个观测点的位置加入均值为 0,方差为 δ 的高斯噪声.学习过程中,系统状态空间离散化为 $60 \times 40 \times 8$ 个状态,即区域按长,宽离散化为 60×40 个网格,将朝向角离散化为 8 个状态,每个状态朝向角跨度 45° ;每间隔 $K = 20$ 个数据采样点进行一次节点选择.视觉信息评价仍采用第 2 节可见性、朝向、清晰度和切换次数的视觉评价准则,其中 A 类场景中目标可见性定义为目标-摄像机连线与该摄像机视域中轴夹角小于一定阈值, B 类场景中目标可见性在夹角判断基础上还需判断目标-摄像机连线与障碍物矩形是否相交;假设部署在场景中的摄像机硬件配置相同,目标成像清晰度在一定程度上可反映为目标与摄像机的位置远近关系,即位于 (x_s, y_s) 的目标相对位置为 (x_c, y_c) 摄像机的清晰度评分可类似于式(4)表示为 $M_c^a = 1 - \sqrt{(x_s - x_c)^2 + (y_s - y_c)^2} / \text{MaxDis}$.

为了分析方法有效性和收敛性,本文选取 A, B 两个场景各 20 000 条运动轨迹作为算法输入,初始 Q 值表分别按第 1.3 节方法(with_initiation)和初始值为 0(without_initiation)进行设置.采用本文方法 A, B 场景(scene_A 和 scene_B)下目标轨迹采样点的单步回报均值分别如图 3(a)、和图 3(b)所示,轨迹中各插曲的平均切换次数如图 3(e)、图 3(f)所示.若在每个决策点上作贪心选择(greedy_selection),即选取当前视觉评分最大的节点作为最优摄像机, A, B 场景下目标轨迹采样点的单步回报均值(考虑切换代价)分别如图 3(c)、图 3(d)所示,各插曲平均切换次数分别如图 3(g)、图 3(h)所示.从图 3 中可看出,采用本文方法, A, B 场景的采样点视觉评分均值与插曲的平均切换次数在轨迹数为 2 000 附近趋向收敛,若 Q 值表初始化设置为 0,则收敛速度较慢,说明采用拓扑信息作为先验知识可有效提高算法收敛速度,可显著减少轨迹训练的数量.同时对比图 3(a)、图 3(b)、图 3(e)、图 3(f)与图 3(c)、图 3(d)、图 3(g)、图 3(h),本文方法在 A, B 场景仿真实验结果的单步回报均值及切换次数方面优于贪心选择结果.

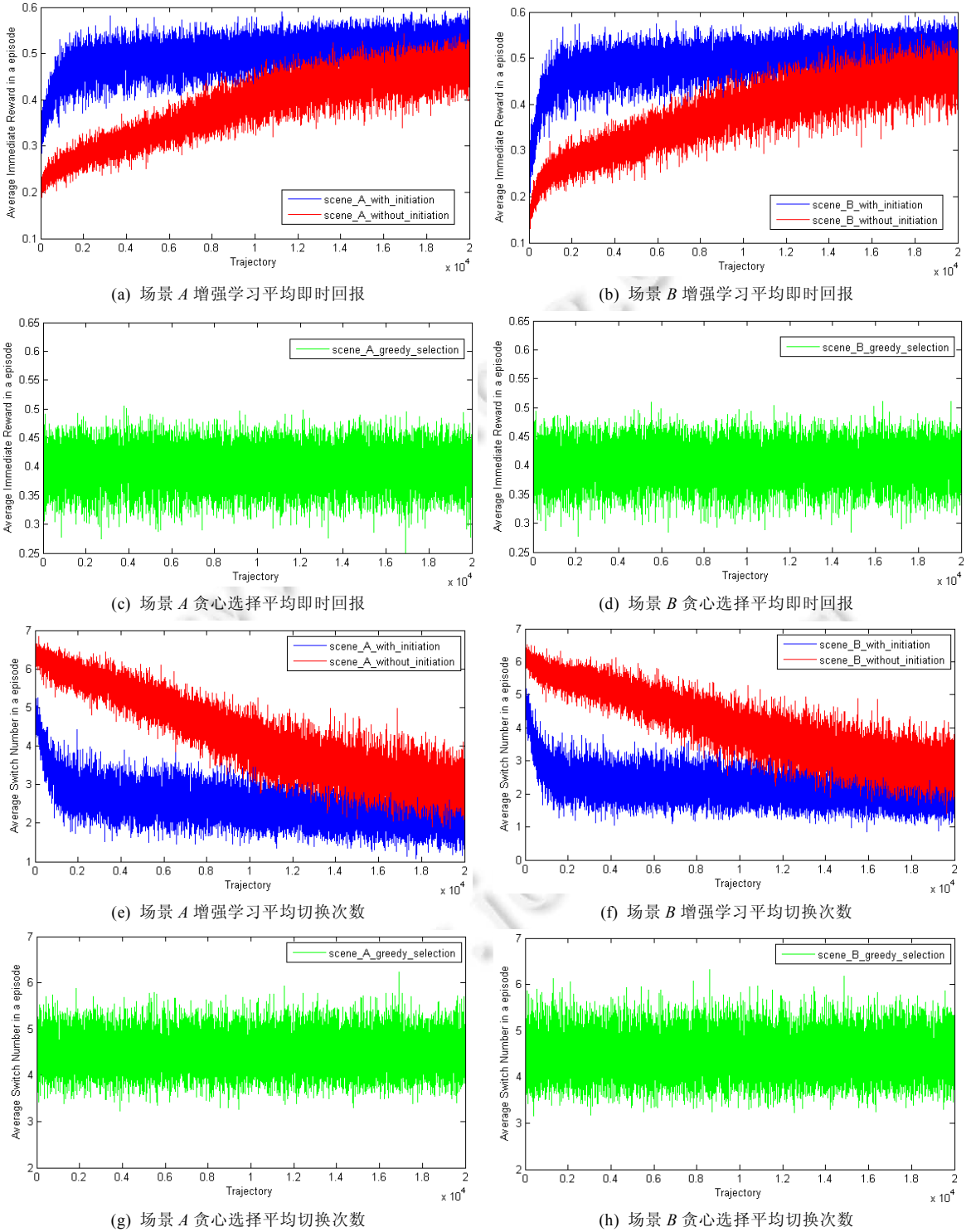


Fig.3 Result of simulation

图 3 虚拟实验结果

3.3 真实场景实验结果与分析

本文在真实场景中部署 12 个已标定的摄像机,录制每组 10 分钟,共 60 组约 10 小时的视频作为训练数据,另选 5 组每组 2 000 帧作为与其他方法对比的测试数据,测试过程不再进行新动作尝试和策略更新.由于真实场景中的目标运动轨迹较仿真环境更为规律,场景中摆放物体占用一定空间,因此实际需要训练的状态空间更小,收敛速度较仿真环境下更快.在测试数据学习过程中采用 VTD 跟踪器^[29]进行目标跟踪,考虑到跟踪算法不是本文考察重点,对跟踪过程中目标丢失情况进行了手工修正处理.测试视频序列 1 采样及目标跟踪结果如图 4 所示,根据跟踪结果,采用式(10)和式(11)分别计算被监控目标相对摄像机 i 朝向评分 M_d^i 和图像清晰度 M_c^i 评分.对比各摄像机的朝向评分可知,当目标正对摄像机运动时具有较高朝向评分,在视频跟踪结果相对稳定的状态下,靠近摄像机可获取较清晰的目标区域图像,清晰度评分也相对较高.

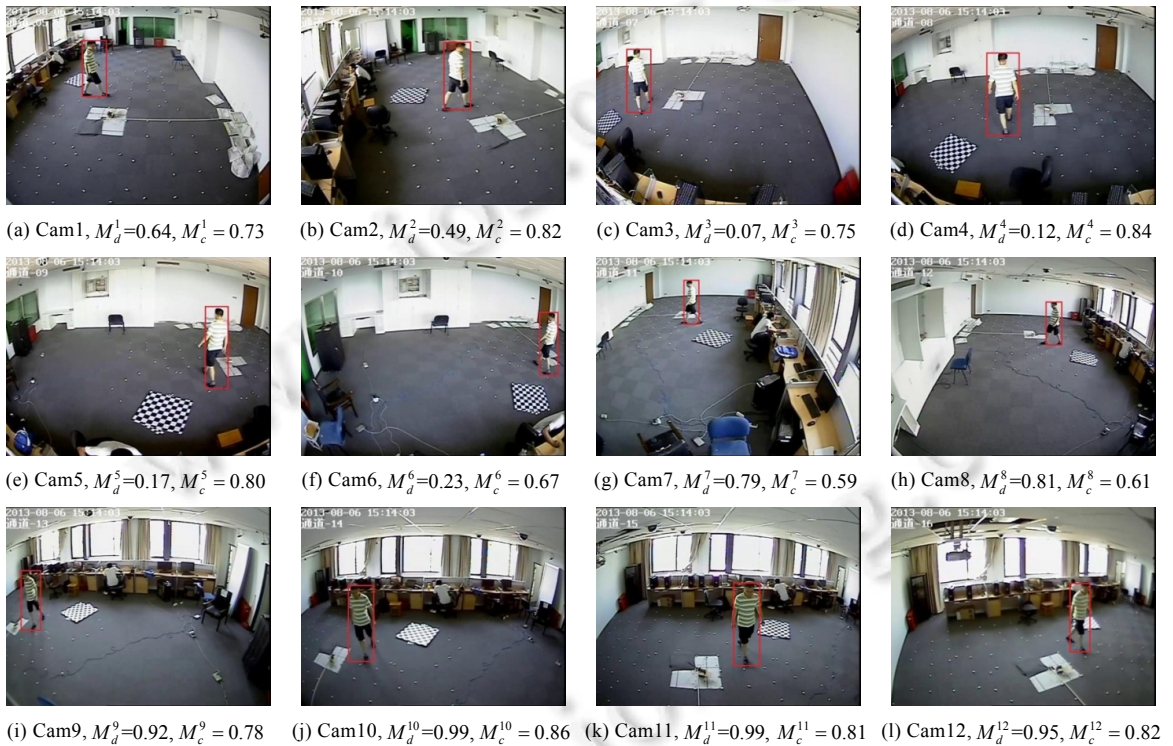
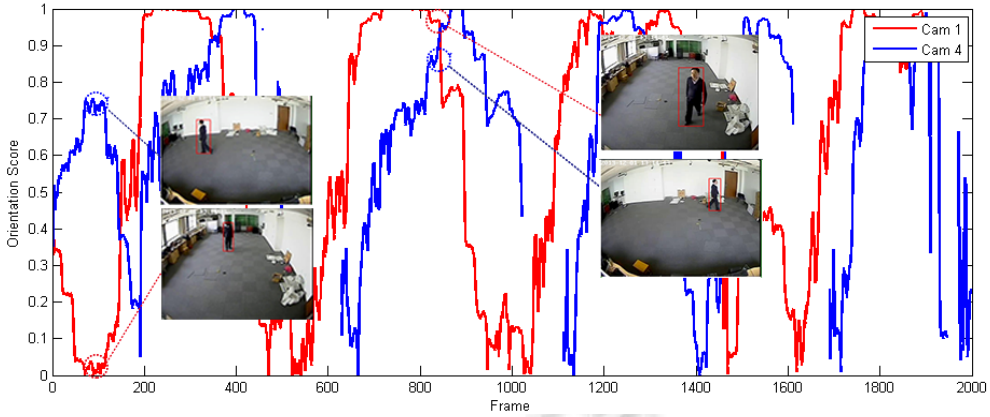


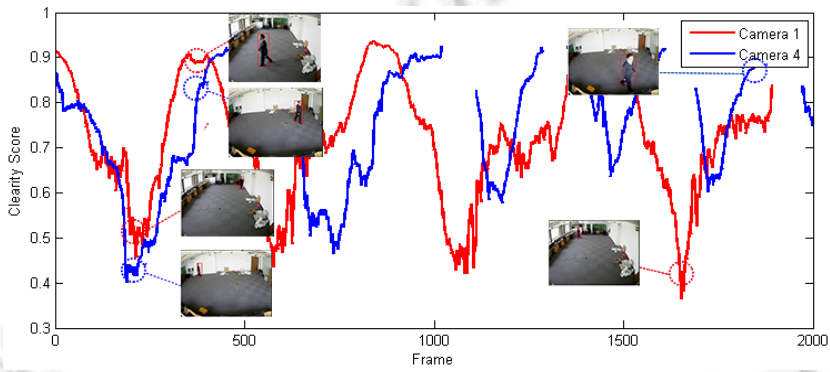
Fig.4 Sample frames and tracking result of video sequence 1

图 4 视频序列 1 采样帧与跟踪分析结果

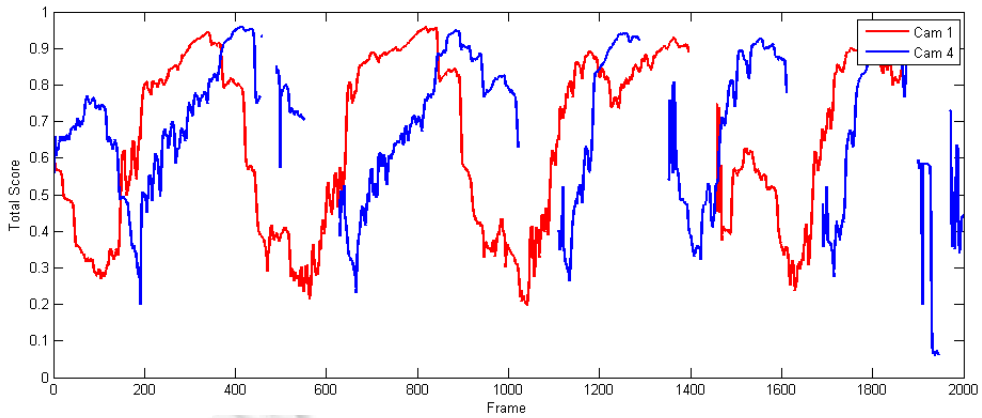
以测试视频序列 2 为例,视角 1(Cam1)和视角 4(Cam4)所计算的朝向、清晰度和总评分曲线分别如图 5(a)~图 5(c)所示.在朝向评分曲线中,第 90 帧和第 830 帧目标相对于 Cam1 和 Cam4 的 x 轴角度真值分别为 0.09 和 2.65,第 90 帧朝向评分分别为 0.007 和 0.75,第 830 帧朝向评分分别为 0.97 和 0.84,均较好地反映了目标相对于不同摄像机的角度信息.当目标重新出现在场景中时,部分朝向评分曲线(如 Cam4 在 1 350 帧和 1 950 帧附近)出现不稳定的现象,也使总评分曲线相应产生了不同程度抖动,这主要是由于跟踪器初始状态下性能不稳定,误差较大,同时可用于拟合直线点数较少造成的.在清晰度评分曲线图中,与评分曲线中对应的第 210 帧、第 380 帧、第 1 650 帧和第 1 870 帧图像分别如图所示,结合曲线和图像可以看出,当清晰度评分较低时,目标距离摄像机较远,目标图像清晰度也较低,反之,目标的视觉辨识度较高,由于目标跟踪算法与标定存在一定的误差以及受到背景的影响,清晰度曲线在部分区域(如 Cam1 的 1 150 帧附近)产生了一定程度的抖动.



(a) Cam1和 Cam4朝向评分曲线



(b) Cam1和 Cam4清晰度评分曲线



(c) Cam1和 Cam4视觉总评分曲线

Fig.5 Curves of visual information score for Cam1 and Cam4

图 5 Cam 1 和 Cam 4 视觉信息评分曲线

我们将本文方法(RL)与 Max, GAME, POMDP 方法在 5 组测试数据上进行了对比,需要说明的是,为进行比较,实验中所有摄像机工作并得到跟踪结果,Max, GAME 和 POMDP 方法中评分值采用我们提出的视觉信息评价方法.图 6 是测试视频序列 2 和序列 3 的选择结果,对视频与选择结果分析可以发现,由于跟踪过程中引入的

部分误差导致部分视角下的视觉评分与目标真实运动状态不相一致,使MAX和GAME方法产生了一定程度的误选,尤其是GAME方法在多个视角的视觉评分(个人效用)相近或低于一定阈值时,选择结果易产生抖动效果,如测试视频序列1的第650帧~第750帧和测试视频序列2的第900帧~第1000帧,均出现了较为频繁切换,而本文方法可根据训练数据学习得到的 Q 值表有效减少误选的发生,减少了频繁切换。

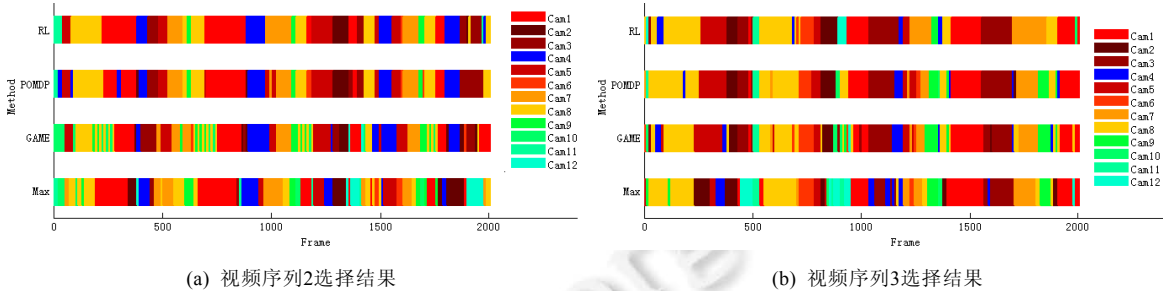


Fig.6 Examples of camera node selection result

图6 摄像机节点选择结果示例

5组视频序列在不同方法下的视频序列中切换总次数(switching number)和单帧平均视觉信息评分均值(visual score per frame)分别如图7(a)、图7(b)所示。从图中可以看出,在单帧平均视觉评分方面,贪心选择在测试的各视频序列中评分较高,但各方法单帧视觉评分总体较为接近,本文方法评分结果接近贪心选择;而在总的切换次数方面,本文方法比其他方法,尤其是贪心选择与博弈论方法有较为明显的优势。

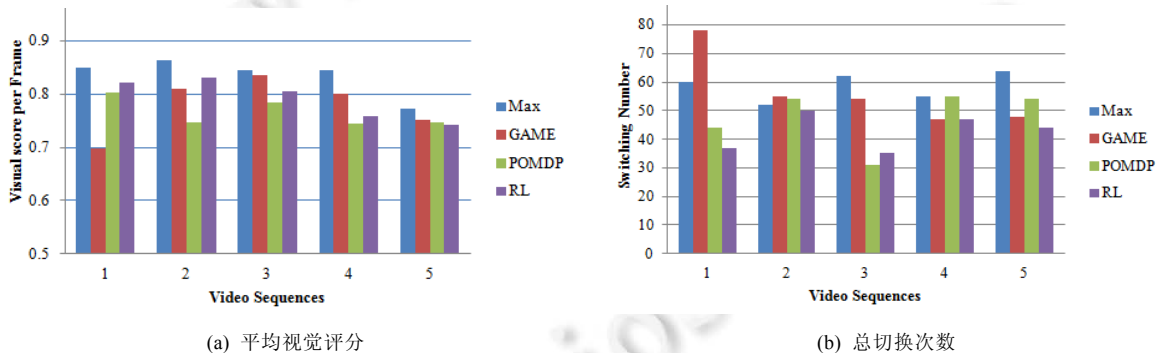


Fig.7 Comparison of visual information score and switch number for different methods

图7 选择方法视觉信息评分与切换次数对比

4 总结

随着摄像机网络广泛应用于安全监控等领域,从海量的实时视频数据中挖掘最富信息量视角成为摄像机网络应用的重点关注的问题之一。为获取监控应用中目标高质量视频,本文提出了一种基于增强学习方法的摄像机网络节点动态选择方法,仿真实验表明,本文提出的摄像机选择方法具有较高的期望回报与较少的视角切换次数。此外,针对真实场景视频数据,本文分析了场景中目标可见性、朝向特征,以及目标区域图像的清晰度,实验结果表明,本文方法可有效地减少因视觉计算误差引起的频繁切换,满足监控应用需求。

本文方法主要研究解决面向单视角期望信息最大化的节点选择与调度问题,在目标数量较少的部分重叠视域场景中取得了较好的效果。在未来的工作中,我们将一方面选择更为稳定的目标检测方法提高视觉信息评价可靠性,另一方面将进一步考察网络带宽和计算能力等条件受限情况下面向多个目标的摄像机网络节点协作监控问题。

References:

- [1] Soro S, Heinzelman W. A survey of visual sensor networks. *Advances in Multimedia*, 2009,1–22. [doi: 10.1155/2009/640386]
- [2] Foresti GL, Micheloni C, Piciarelli C, Snidaro L. Visual sensor technology for advanced surveillance systems: Historical view, technological aspects and research activities. *Sensors*, 2009,9(4):2252–2270. [doi: 10.3390/s90402252]
- [3] Charfi Y, Wakamiya N, Murata M. Challenging issues in visual sensor networks. *IEEE Wireless Communications*, 2009,16(2):44–49. [doi: 10.1109/MWC.2009.4907559]
- [4] Ercan AO, Gamal AE, Guibas LJ. Object tracking in the presence of occlusions via a camera network. In: *Proc. of the 6th Int'l Conf. on Information processing in sensor networks*. IEEE Press, 2007. 509–518. [doi: 10.1109/IPSNS.2007.4379711]
- [5] Han B, Joo SW, Davis LS. Multi-Camera tracking with adaptive resource allocation. *Int'l Journal of Computer Vision*, 2011,91(1): 45–58. [doi: 10.1007/s11263-010-0373-3]
- [6] Li W, Zhang W. Sensor selection for improving accuracy of target localization in wireless visual sensor networks. *IET Wireless Sensor Systems*, 2012,2(4):293–301. [doi: 10.1049/iet-wss.2012.0033]
- [7] Soro S, Heinzelman W. Camera selection in visual sensor networks. In: *Proc. of the IEEE Conf. on Advanced Video and Signal Based Surveillance*. IEEE Press, 2007. 81–86. [doi: 10.1109/AVSS.2007.4425290]
- [8] Tessens L, Morbee M, Lee H, Philips W. Principal view determination for camera selection in distributed smart camera networks. In: *Proc. of the 2nd ACM/IEEE Int'l Conf. on Distributed Smart Cameras*. IEEE Press, 2008. 1–10. [doi: 10.1109/ICDSC.2008.4635699]
- [9] Rudoy D, Lih Z. Viewpoint selection for human actions. *Int'l Journal of Computer Vision*, 2011,97(3):243–254. [doi: 10.1007/s11263-011-0484-5]
- [10] Fu Y, Guo Y, Zhu Y. Multi-View video summarization. *IEEE Trans. on Multimedia*, 2010,12(7):717–729. [doi: 10.1109/TMM.2010.2052025]
- [11] Mo Y, Ambrosino R, Sinopoli B. Sensor selection strategies for state estimation in energy constrained wireless sensor networks. *Automatica*, 2011,47(7):1330–1338. [doi: 10.1016/j.automatica.2011.02.001]
- [12] Feixas M, Sbert M, Gonzalez F. A unified information-theoretic framework for view-point selection and mesh saliency. *ACM Trans. on Graphics*, 2009,6(1):1–23. [doi: 10.1145/1462055.1462056]
- [13] Leifman G, Shtrom E, Tal A. Surface regions of interest for viewpoint selection. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE Press, 2012. 414–421. [doi: 10.1109/CVPR.2012.6247703]
- [14] Park J, Bhat C, Kak AC. A look-up table based approach for solving the camera selection problem in large camera networks. In: *Proc. of the Int'l Workshop on Distributed Smart Cameras*. ACM Press, 2006. [doi: 10.1.1.144.2859]
- [15] Monari E, Kroschel KA. Knowledge-Based camera selection approach for object tracking in large sensor networks. In: *Proc. of the 3rd ACM/IEEE Int'l Conf. on Distributed Smart Cameras*. IEEE Press, 2009. 1–8. [doi: 10.1109/ICDSC.2009.5289400]
- [16] Bimbo AD, Pernici F. Towards on-line saccade planning for high-resolution image sensing. *Pattern Recognition Letters*, 2006, 27(15):1826–1834. [doi: 10.1016/j.patrec.2006.02.014]
- [17] Li Y, Bhanu B. Utility-Based camera assignment in a video network: A game theoretic framework. *IEEE Sensors Journal*, 2009, 11(3):676–687. [doi: 10.1109/JSEN.2010.2051148]
- [18] Jiang H, Fels S, Little J. Optimizing multiple object tracking and best view video synthesis. *IEEE Trans. on Multimedia*, 2008, 10(6):997–1012. [doi: 10.1109/TMM.2008.2001379]
- [19] Daniyal F, Taj M, Cavallaro A. A content and task-based view selection from multiple video streams. *Multimedia Tools and Applications*, 2010,46(2):235–258. [doi: 10.1007/s11042-009-0355-z]
- [20] Krishnamurthy V, Djonin DV. Structured threshold policies for dynamic sensor scheduling—A partially observed Markov decision process approach. *IEEE Trans. on Signal Processing*, 2007,55(10):4938–4957. [doi: 10.1109/TSP.2007.897908]
- [21] Spaan MTJ, Lima PU. A decision-theoretic approach to dynamic sensor selection in camera networks. In: *Proc. of the Int'l Conf. on Automated Planning and Scheduling*. AAAI Press, 2009. 1–8.
- [22] Li Q, Sun Z, Chen S, Liu Y. A method of camera selection based on partially observable Markov decision process model in camera networks. In: *Proc. of the American Control Conf*. IEEE Press, 2013. 3833–3839.

- [23] Kaelbling RE, Littman ML, Cassandra AR. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 1998,101(1):99–134.
- [24] Gupta A, Mittal A, Davis LS. Cost: An approach for camera selection and multi-object inference ordering in dynamic scenes. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*. IEEE Press, 2007. 1–8. [doi: 10.1109/ICCV.2007.4408842]
- [25] Shen C, Zhang C, Fels S. A multi-camera surveillance system that estimates quality-of-view measurement. In: *Proc. of the IEEE Int'l Conf. on Image Processing*. IEEE Press, 2007. 193–196. [doi: 10.1109/ICIP.2007.4379279]
- [26] Agarwal A, Triggs B. Recovering 3D human pose from monocular images. *IEEE Trans. on Pattern analysis and Machine Intelligence*, 2006,28(1):44–58. [doi: 10.1109/TPAMI.2006.21]
- [27] Watkins C, Dayan P. *Q-Learning*. *Machine Learning*, 1992,8(3):279–292.
- [28] Sutton R, Barto A. *Reinforcement Learning: An Introduction*. MIT Press, 1999.
- [29] Kwon J, Lee K. Visual tracking decomposition. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. San Francisco: IEEE Press. 2010. 1269–1276. [doi: 10.1109/CVPR.2010.5539821]



李骞(1980—),男,湖南湘潭人,博士生,主要研究领域为计算机视觉,人工智能.



陈松乐(1976—),男,博士,主要研究领域为模式识别.



孙正兴(1964—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为多媒体计算,计算机视觉,智能人机交互.



夏士明(1990—),男,硕士生,主要研究领域为计算机视觉.