

# 家庭智能空间下基于场景的人的行为理解

田国会, 吉艳青, 李晓磊

(山东大学 控制科学与工程学院, 山东 济南 250061)

**摘要:**为了更好地在日常生活中给人提供智能化服务,对家庭环境下人的行为理解问题进行了研究.首先利用运动目标检测方法提取运动人体在环境中的坐标,然后结合行为特点把场景划分成不同区域,建立人体在环境中的位置关联矩阵和时空关联矩阵.通过马尔可夫模型统计出人体在空间中的位置状态转移概率矩阵及其状态持续时间矩阵,生成日常行为模板.根据当前行为与日常行为模板的相似度可检测出反常习惯和突发异常行为,同时可根据不同区域的行为模式分析人的意图.在智能空间平台下利用机器视觉技术基于场景信息实现了人的行为理解,并通过实验表明了方法的有效性.

**关键词:**智能空间;机器视觉;场景信息;行为理解;异常检测;意图识别

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 1673-4785(2010)01-0057-06

## Human behaviors understanding based on scene knowledge in home intelligent space

TIAN Guo-hui, JI Yan-qing, LI Xiao-lei

(School of Control Science and Engineering, Shandong University, Ji'nan 250061, China)

**Abstract:** Challenges in understanding human behavior in a home environment were studied in order to provide more intelligent services. First, spatial coordinates of human bodies were extracted from motion detector data. Then, by dividing the environment into different stations and observing the types of behavior typical at various periods of time in that area, a station-based occupational matrix with a time dimension was established. After establishing the station state transitional probability matrix and the state duration time distribution matrix based on the Markov model, a daily behavioral template was constructed. Behavior outside of normal habits as well as behavior resulting from unexpected accidents could be detected in real-time by comparing the similarity of current behavior with templates showing typical daily behavior. At the same time, human intention could be predicted based on behavior patterns typical in different areas. In this way, better understanding of human behavior becomes possible. The effectiveness of this method was proved by experiments.

**Keywords:** intelligent space; machine vision; scene knowledge; behavioral understanding; anomaly detection; intention recognition

家庭服务机器人<sup>[1]</sup>设计的主要目标之一就是能通过观察人的行为,正确理解人的意图,对病人、老人和小孩等进行监护.因此,行为理解是服务机器人作为人提供服务的基础.现有的大部分行为理解方法都是基于单一的机器视觉<sup>[2-4]</sup>.其方法主要有3类:第1类,基于时空特征的模式分析方法<sup>[5-6]</sup>,该方法首先将提取人体区域图像序列转换为一组静态形状模式,然后在识别阶段和预先存储的行为样本相

比较.这类方法计算复杂度低、实现简单,但是对噪声和运动时间间隔的变化敏感.第2类,基于模型的方法<sup>[7-8]</sup>,该方法先采用人体2D或3D模型得到每一帧视频中的人体姿势,然后利用随时间变化的姿势来描述不同动作.该方法准确度高,但是计算量大,对设备要求高,实时性差.第3类,基于图像统计处理的方法<sup>[9-10]</sup>,主要是提取动作的关键帧图像特征,进行训练和识别.这类方法能很好地识别固定的姿态或者简单的动作.

家庭环境下拍摄的图像背景杂乱,人体姿态复杂,衣着颜色风格各异,运动的时间间隔变化多端,因

收稿日期:2008-11-22.

基金项目:国家高技术研究发展计划重点资助项目(2006AA040206).

通信作者:田国会. E-mail: g. h. tian@sdu.edu.cn.

此前面描述的行为分析的方法都很难应用于家庭行为理解系统. 本文避开家庭环境中人体姿态提取难的问题和当前行为分析方法只对单个姿态或者一组动作序列理解的局限, 结合家庭智能空间平台提供的环境数据和区域行为模式、日常行为习惯信息, 利用马尔可夫模型建立行为模板, 实现了人的反常习惯、突发异常行为的实时检测和人意图的正确理解.

## 1 行为理解方案

家庭服务机器人行为理解系统需要识别人的意图, 预测人在行为过程的需要, 提供适当的服务以提高人的生活质量; 检测人的异常行为(反常习惯和突发异常行为), 监控人的健康状况, 提高人生活的安全感. 涉及的关键问题是要建立一种适合机器人提供自主服务的意图识别模型; 并且建立一种适合服务机器人自主学习服务对象行为习惯和发现异常行为的认知模型. 家庭环境下的行为理解系统具体实现步骤如下:

1) 首先阐述智能空间技术平台和智能空间信息系统中基于场景的区域行为模式和行为习惯.

2) 行为描述. 利用多个摄像机的运动人体检测方法确定人体在家庭环境中的位置, 根据智能空间中场景的区域划分, 将视频图像中获得的人体位置关联到不同的区域, 同时记录关联的顺序和时间, 构建位置关联矩阵和时空关联矩阵; 统计出人体在空间中的位置状态转移概率和位置状态持续时间概率, 生成日常行为模板.

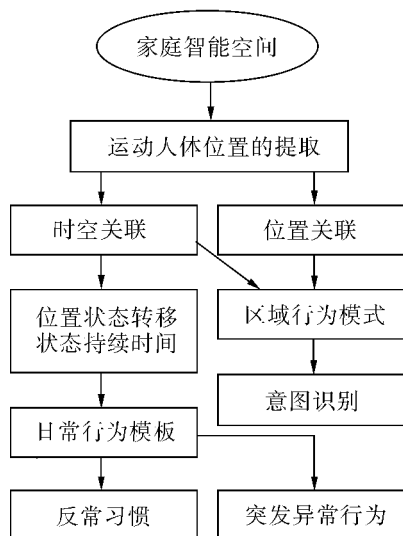


图1 系统流程图

Fig. 1 The system flow chart

3) 行为检测. 将当前的行为数据与日常行为模板匹配, 根据匹配的相似度, 判断人的反常习惯、突发异常行为; 通过匹配人的位置、时间和区域行为模

式来分析人的意图.

系统框图见图1, 下文将具体论述各个部分的实现过程, 首先描述家庭智能空间的技术平台以及本文所利用的信息.

## 2 基于场景的行为模式和行为习惯

### 2.1 智能空间技术平台

应用于家庭服务机器人的智能空间把感知器件、执行器件分布在空间中相应的位置, 实现对其中物的位置、形状以及人的动作、行为等全面的感知和理解, 通过执行器实现相应的服务任务, 同时能够为机器人提供更加完备的环境信息, 实现更加快捷、准确、稳定的家庭服务工作<sup>[11]</sup>. 本文的家庭智能空间中主要组成部分包括: 分布式摄像机、网络设备、数据库和信息处理系统. 图2所示是一个典型的家庭智能空间的环境, 分别在客厅、卧室、书房安装分布式摄像机, 分布式摄像机的安装要尽可能保证多视角、全方位地获取场景视频信息.

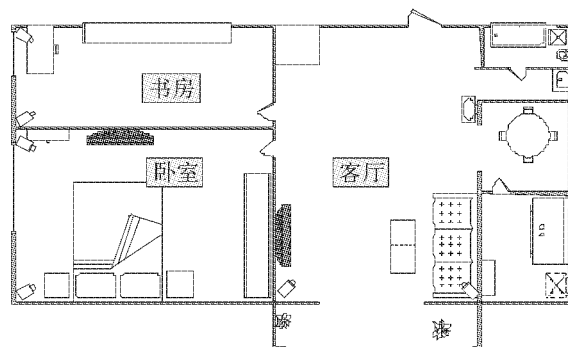


图2 家庭智能空间的典型环境

Fig. 2 Home intelligent space

### 2.2 行为模式和行为习惯的数学表示

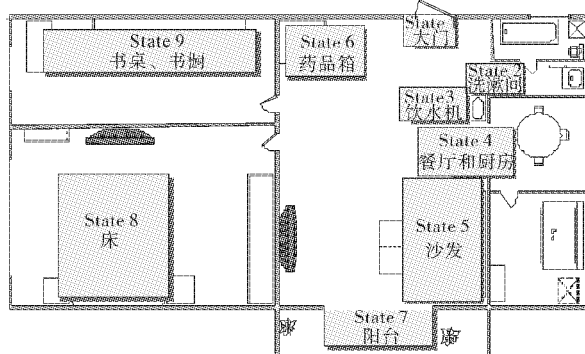


图3 家庭智能空间区域状态模型

Fig. 3 Home intelligent space scene state model

不同的场景对应不同的行为, 例如, 人在书桌和书橱前对应的正常行为是读书, 突发异常行为是长时间地趴在书桌上. 根据不同场景对应不同行为模式, 将家庭智能空间划分为不同的区域状态. 图3表示数据库系统中存储的基于场景的区域状态划分,

共 10 个状态. State1 ~ State9 的区域如图 3 所示,其中 State0 是去除已经划分的状态外的区域,描述为自由区域状态.

基于场景的行为模式是统计人体在不同区域状态的活动,将其划分为区域日常行为模式和突发异常行为模式. 区域的日常行为模式,是人体在该区域中经常从事的正常的行为,区域的日常行为模式是下文中的人的意图理解的基础. 表 1 描述了根据家庭智能空间的区域行为模式,不同区域的日常行为模式各不相同,例如沙发区域(State5)的日常行为模

式是看电视和休息(Rest/Watch TV),书桌区域 State9 的日常行为模式为学习(Study).

把老年人在日常生活中出现的非主观意愿支配的,具有一定特殊性、偶然性、不可以准确预期的异于正常行为的情况统称为异常行为. 异常行为包括突发异常行为和反常习惯. 不同区域的突发异常行为模式是相同的,主要有病倒事件和长时间静止(Fall down with an illness,keep still for long time),如表 1.

表 1 基于场景的行为模式划分  
Table1 Patterns of behavior based on the scene

模式划分	区域状态									
	State0	State1	State2	State3	State4	State5	State6	State7	State8	State9
日常行为模式	Walk/Stand	Out/In	Rinse	Drink	Eat	Rest/TV	Take medicine	Enjoy	Sleep	Study
突发异常行为模式	Fall down with an illness,keep still for long time									

本文根据人在房间中不同区域状态的活动顺序来描述人的行为习惯. 例如,早晨人的正常行为习惯顺序是:起床——洗漱——喝水——吃饭——看电视——出门锻炼;如果当前行为的状态转移顺序为起床——洗漱——喝水——吃饭——看新闻——休息,则被检测为反常习惯. 表 2 描述了时段  $T_1$  (7:00—9:00) 一个日常行为习惯的例子,人的状态转移的顺序是 State8 - State2 - State3 - State4 - State5 - State1, 状态持续的时间分别是 32 min、11 min、4.83 min、20.67 min、32 min、16 min. 其中, State0 不参与人的日常行为习惯的描述,因此表 2 中没有标记 State0 持续的时间.

表 2 家庭智能空间中一个日常行为习惯的例子  
Table 2 The example of human daily behavior in home intelligent space

状态区域	起始时间	持续时间/min
State8	7:00—7:32	32.00
State2	7:33—7:44	11.00
State3	7:45—7:49:50	4.83
State4	7:50—8:10:40	20.67
State5	8:11—8:43	32.00
State1	8:44—9:00	16.00

3 行为描述及检测

运动人体的检测是行为理解的基础,利用多个摄像机检测人体运动,同时确定人体在环境中的位置<sup>[12]</sup>. 然而,单纯地确定人体在环境中坐标、持续时

间,对行为理解没有太大的意义,只有结合周围的场景以及一些历史数据才能更好地理解人的行为. 下文将详细阐述行为的描述方法和检测方法.

3.1 位置关联矩阵的建立

位置关联矩阵是对图像中当前人体位置的观测值和场景中物品区域进行关联而建立的矩阵. 首先通过运动检测提取运动人体在图像中的位置,同时通过坐标转换和身高模型确定人体在环境中的坐标<sup>[12]</sup>,然后利用已经划分的房间的区域状态,判断人体位置和区域的相对关系建立矩阵,最后结合区域状态行为模式分析人的意图. 在位置关联过程中需要注明的是:若运动人体检测超出视觉可见区域,则定义人体依然滞留在前一个区域状态.

位置关联矩阵记为  $R$ ,建立过程如下:首先以场景中物品所占的区域和当前所检测目标的环境坐标为元素建立矩阵  $D$ ,“行”为当前帧观测目标  $O_{(x,y)}$ ,  $x,y$  为目标在环境中的坐标值;“列”为物品区域  $\{S_1, S_2, \dots, S_n\}$ ,  $n$  为已划分的物品区域的数量. 基于  $D$  矩阵,建立判断矩阵  $R$ ,对当前帧的运动人体目标质心  $O_{(x,y)}$ ,判断与物品区域  $S_i$  的关系,若  $O_{(x,y)} \in S_i$  则置  $R$  中相应值为 1,否则为 0. 人体位置与 State5 关联的实例,如表 3.

表 3 位置关联矩阵  $R$   
Table 3 Location matrix  $R$

区域状态	人体位置	区域状态	人体位置
state0	0	state5	1
state1	0	state6	0

续表3

区域状态	人体位置	区域状态	人体位置
state2	0	state7	0
state3	0	state8	0
state4	0	state9	0

### 3.2 时空关联矩阵

人体行为不仅表现在位置的关联,还表现在位置状态持续时间和转移到概率。例如,人早上醒来是从床的位置转移到沙发位置还是转移到洗漱间位置?在沙发区域看5分钟电视还是2个小时?本文以时空关联矩阵记录人体在不同区域活动的持续时

间和顺序关系来表征人的行为。

时空关联矩阵记为  $T$ , 以运动人体在场景中的区域位置、所属时间段、持续时间为元素。“列”为当前人体所属的区域  $S_i$ , 是一个动态增长的序列;“行”为划分的时段  $\{T_1, T_2, \dots, T_x\} (x=1, 2, \dots, 12)$  以及人体在区域  $S_i$  的持续时间  $\tau^{S_i}$ , 其中  $\tau^{S_i}$  是人体从区域  $S_i$  转移到区域  $S_j$  的时间间隔, 如式(1)。

$$\tau^{S_i} = T(S_i) - T(S_j). \quad (1)$$

以某天时段  $T_1$  人体在家庭环境中的行为为例建立时空关联矩阵如表4。

表4 时空关联矩阵  $T$ Table 4 Time and location matrix  $T$ 

min

时间参数	区域状态									
	State8	State0	State2	State3	State0	State4	State0	State5	State0	State1
时段	$T_1$	$T_1$	$T_1$	$T_1$	$T_1$	$T_1$	$T_1$	$T_1$	$T_1$	$T_1$
持续时间	32	1	11	4.9	0.1	20.8	0.2	32	1	16

### 3.3 位置状态转移概率

人体在不同区域之间的活动可以看做是在不同状态之间的转移,利用马尔可夫模型建立人体在智能空间中的位置状态转移矩阵  $F = [f_{ij}]$ 。链的状态空间记做  $I = \{S_1, S_2, \dots, S_n\}$ , 统计人体位置从  $S_i$  区域关联到  $S_j$  的频率记做  $f_{ij}$ , 归一化  $f_{ij}$  获得状态转移概率  $m_{ij}$ , 根据  $m_{ij}$  建立位置状态转移概率矩阵  $M$ , 如式(2)。

$$M = [m_{ij}] = \left[ \frac{f_{ij}}{\sum_{i,j} f_{ij}} \right]. \quad (2)$$

根据智能空间数据库系统中记录的人的日常行为习惯,利用位置状态转移概率建立行为模板  $M_{\text{daily}}$ 。时段  $T_1$  的  $M_{\text{daily}}$  图示形式如图4, 其中线宽代表转移概率的大小, 很小的概率没有表示。

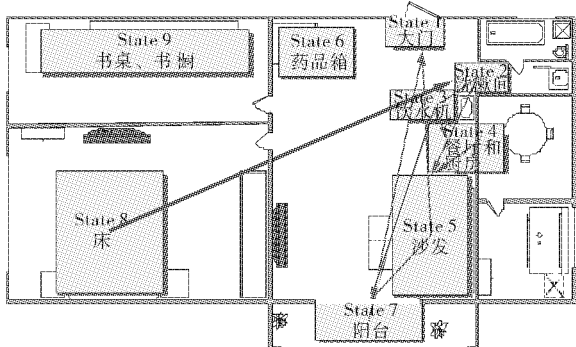


图4 位置状态转移模板

Fig. 4 The station state transition template

### 3.4 位置状态持续时间

马尔可夫链中人体在不同状态的状态持续时间

为  $\tau^{S_i}$ , 利用高斯函数建立数学模型。统计  $N$  天中时段  $T_x$  人体在区域  $S_i$  的状态持续时间  $\tau^{S_i}$ , 建立高斯函数  $P(\tau^{S_i}) = f(T_x, S_i, \tau, u, \sigma)$ , 其中  $u, \sigma$  表示人体关联到区域  $S_i$  持续时间的均值和方差。人在不同状态持续时间的习惯矩阵为

$$H_{\text{daily}} = \left[ \sum_{i=1}^n P(\tau^{S_i}) \right].$$

### 3.5 行为检测器

本文将异常行为分为反常习惯和突发异常行为, 分别设计检测器对其检测, 同时利用位置关联和时空关联的时段信息分析人的意图。

#### 3.5.1 反常习惯的检测

时空关联矩阵获得人体在智能空间中的一组状态  $S^k (k=1, \dots, n)$ , 计算状态序列  $S^k$  的状态转移与日常行为模板  $M_{\text{daily}}$  的相似度  $E_S^{\text{likelihood}}$ 。

$$E_S^{\text{likelihood}} = \frac{1}{n-1} \sum_{k=1}^n m_{S^1:n}. \quad (3)$$

如果  $E_S^{\text{likelihood}}$  的数值低于一定的阈值, 则认为是反常习惯。这种方法能有效地检测反常习惯, 但是对于摔倒、静止、病倒等突发的异常行为不能实时地识别, 因而需要设计突发的异常行为检测器。

#### 3.5.2 突发异常行为的检测

实时地计算状态  $S^k$  持续时间与持续时间习惯矩阵  $H_{\text{daily}}$  的相似度  $E_\tau^{\text{likelihood}}$ 。

$$E_\tau^{\text{likelihood}} = P(\tau^{S^k}). \quad (4)$$

如果  $E_\tau^{\text{likelihood}}$  的数值低于一定的阈值 (本文设定为 0.3) 且状态  $S^k$  的持续时间大于该状态的持续

时间均值,则识别为突发异常行为.为了减少不必要的计算量,每隔一分钟计算一次状态  $S^k$  的持续时间相似度  $E_{\tau}^{\text{likelihood}}$ .

如果人的位置关联到某个区域状态,长时间内运动输出为0,即人体处于长时间静止状态,则认为是严重的突发异常情况.

### 3.5.3 意图分析

当未检测到人的异常行为时,根据位置关联矩阵、时空关联矩阵和场景信息分析人的意图,进行相应的服务.例如,当人在床的位置(State8),时空关联矩阵的时段是傍晚时段( $T_7$ ),则意图分析的结果是人要休息,启动的服务则是关闭门窗.

## 4 实验结果与分析

### 4.1 反常习惯检测

根据智能空间信息系统用户的行为数据建立  $T_1$  时段的行为模板  $M_{\text{daily}}^{T_1}$ .统计连续24天  $T_1$  时段的行为数据用于人的行为检测,计算得到行为模板相似度  $E_S^{\text{likelihood}}$ ,如图5,其中纵坐标表示状态转移相似度,横坐标表示时间.

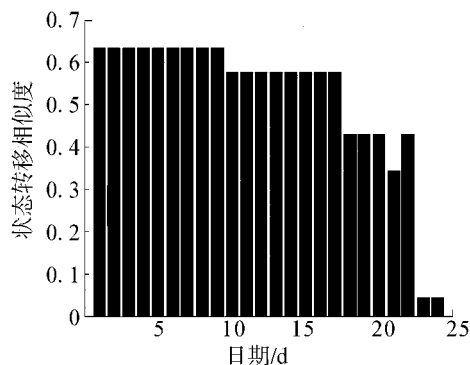


图5 反常习惯检测

Fig. 5 Abnormal habit detection

从图中可以看出第23、24天状态转移相似度的数值很低,被检测为反常的习惯.查看智能空间中数据库中记录的信息发现,第23天和24天人处于生病状态,频繁地在饮水机和药箱2个状态之间转移,而日常行为模板中饮水机和药箱2个状态的转移概率很小,因而状态转移相似度很小.实验表明  $M_{\text{daily}}$  能够有效地检测出人的反常行为习惯.

### 4.2 突发异常行为检测

同样以上述的用户习惯为例,检测人的突发异常行为.已知  $T_1$  时段中的  $M_{\text{daily}}^{T_1}$ 、 $H_{\text{daily}}^{T_1}$  和某天时空关联的状态序列  $S^k (k=1, \dots, n)$ ,状态的持续时间相似度如图6所示.

7:38分首次采集到人在状态 State2 的持续时间相似度小于0.3且大于持续时间均值——预警,

7:42分第二次预警,7:46分报警.7:38—7:46时间段已经有2次预警,但是没有接收到人的语音销警,服务机器人根据突发异常行为处理模式通过网络通知监护人.7:48分时检测到状态持续时间相似度大于0.3,人的状态恢复正常.

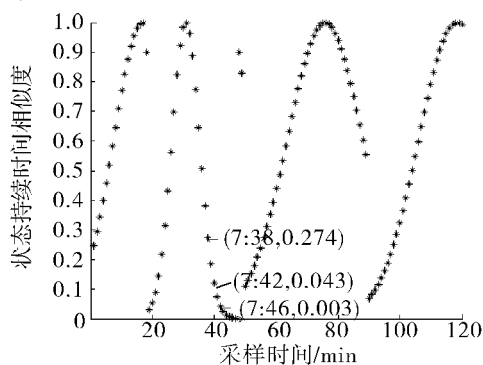


图6 突发异常行为检测

Fig. 6 Abnormal accident behavior detection

### 4.3 意图识别

意图识别是为了预测人在行为过程的需要,提供适当的服务,因而这种服务的提供具有主观性因素和个性化特征.本文设计了一个意图识别及提供服务的示例:即检测到人的状态关联到书桌时,结合光照度传感器信息,自动调亮房间的光源;当检测到人离开书桌时,自动调暗房间光源.

## 5 结束语

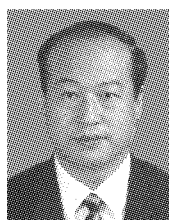
提出了一种新的利用智能空间信息,基于场景的行为理解方法,建立了一种适合机器人提供自主服务的意图识别模型和发现异常行为的认知模型.本文行为理解方法的优点是不需要准确地分割人体各个部分提取人体的姿态,计算复杂度小,能够实时地识别突发异常行为、反常习惯、理解人的意图.不足之处是没有结合时段之间上下文的关系,有待进一步发现习惯数据中所包括的一些深层次的信息.实验表明利用视觉技术结合智能空间信息的行为理解方法具有可行性,可实现独居老人的实时监护,提高了人们的生活质量和安全性.

## 参考文献:

- [1] 田国会. 家庭服务机器人研究前景广阔[J]. 国际学术动态, 2007(1): 28-29.  
TIAN Guohui. Wide future for home service robot research [J]. International Academic Developments, 2007(1): 28-29.
- [2] WANG Liang, SUTER D. Learning and matching of dynamic shape manifolds for human action recognition[J]. IEEE Transactions on Image Processing, 2007, 16(6): 1646-

- 1661.
- [3] ISMAIL H, DAVID H, LARRY S, et al. Real-time surveillance of people and their activities[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 809-830.
- [4] FERRYMAN J, BORG M, THIRDE D. Automated scene understanding for airport aprons[C]//Proceedings of the 18th Australian Joint Conference on Artificial Intelligence. Sidney, 2005: 593-603.
- [5] BOBICK A F, DAVIS J W. The recognition of human movement using temporal templates[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(3): 257-267.
- [6] KOJIMA A. Generating natural language description of human behaviors from video images[C]//IEEE International Conference on Pattern Recognition. Barcelona, 2000: 728-731.
- [7] LEE M W. A model-based approach for estimating human 3D poses in static images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(6): 905-916.
- [8] CHEUNG G K M, KANADE T, BOUGUET J Y, et al. A real time system for robust 3D voxel reconstruction of human motions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head, USA, 2000: 714-720.
- [9] WATANABE K, IZUMI K, KAMOHARA K, et al. Feature extractions for estimating human behaviors via a binocular vision head[C]//Proceedings of International Conference on Control, Automation and Systems. Seoul, Korea, 2007: 634-640.
- [10] YAMATO J, OHYA J, ISHII K. Recognizing human action in time-sequential images using hidded Markov model[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Champaign, USA, 1992: 379-385.
- [11] 田国会, 李晓磊, 赵守鹏, 等. 家庭服务机器人智能空间技术研究与进展[J]. 山东大学学报:工学版, 2007, 37(5): 53-59.
- TIAN Guohui, LI Xiaolei, ZHAO Shoupeng, et al. Research and development of intelligent space technology for home service robot[J]. Journal of Shandong University: Engineering Science, 2007, 37(5): 53-59.
- [12] FLEURET F, BERCLAZ J, LENGAGNE R, et al. Multi-camera people tracking with a probabilistic occupancy map[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 267-282.

#### 作者简介:



田国会,男,1969年生,教授、博士生导师、工学博士,山东大学控制科学与工程学院副院长,中国人工智能学会理事. 主要研究方向为服务机器人、智能空间、多机器人系统的协调与协作等. 作为课题负责人或主要参加人员已完成包括国家自然科学基金项目、国家 863 计划项目、国防预研项目、中国博士后科学基金项目、山东省自然科学基金项目等 15 项. 获山东省科技进步二等奖、山东省教委科技进步(自然科学理论)一等奖各 1 项. 发表学术论文 110 余篇.



吉艳青,女,1984年生,硕士研究生. 主要研究方向为服务机器人、智能空间、基于视觉的监控和行为理解.



李晓磊,男,1973年生,副教授、硕士生导师、工学博士. 主要研究方向为智能空间技术、复杂系统建模与智能优化算法等. 参与国家及省部级科研项目 10 余项,获省科技进步二等奖 1 项,发表学术论文 40 余篇.