Journal of Southeast University (English Edition) Vol. 24, No. 3, pp. 347 – 350 Sept. 2008 ISSN 1003—7985

# Detecting abnormalities for empty nest elder in smart monitoring system

Yang Lei[1] Yang Luming[1] Man Junfeng[1,2] Liu Guangbin[2]

([1]College of Information Science and Technology, Central South University, Changsha 410083, China)
([2]College of Computer and Communication, Hunan University of Technology, Zhuzhou 412008, China)

**Abstract:** In order to implement the real-time detection of abnormality of elder and devices in an empty nest home, multi-modal joint sensors are used to collect discrete action sequences of behavior, and the improved hierarchical hidden Markov model is adopted to abstract these discrete action sequences captured by multi-modal joint sensors into an occupant's high-level behavior—event, then structure representation models of occupant normality are modeled from large amounts of spatio-temporal data. These models are used as classifiers of normality to detect an occupant's abnormal behavior. In order to express context information needed by reasoning and detection, multi-media ontology (MMO) is designed to annotate and reason about the media information in the smart monitoring system. A pessimistic emotion model (PEM) is improved to analyze multi-interleaving events of multi-active devices in the home. Experiments demonstrate that the PEM can enhance the accuracy and reliability for detecting active devices when these devices are in blind regions or are occlusive. The above approach has good performance in detecting abnormalities involving occupants and devices in a real-time way.

**Key words:** multi-media ontology; semantic annotation; abnormality detection; hierarchical hidden Markov model; pessimistic emotion model

In recent years, with the development of the economy, the empty nest family has become the main habitation fashion of the elderly in China. In order to ease the elderly's lonely feeling and enhance the quality of life, a new model of a smart monitoring system for the empty nest elderly (SMS4ENE) is designed and implemented. This unsupervised system can detect abnormal behaviors of the elderly and devices in a home environment in a real-time way, and take correct measures according to the level and reliability of the abnormalities.

The essential significance of SMS4ENE is that it can perceive diversified environmental states in the home and judge abnormalities such as human reactions. There are two issues to address: one is to recognize high-level event of human, the other is to discover and forecast whether the conditions of active devices in the home are normal.

The key issue of abnormality detection is behavior pattern learning concerning the elderly. Former modeling of the activity patterns of people generally use the hidden Markov model (HMM), Kijak et al. [1] used multiple layers of the HMM to analyze scenes in sport videos. Generally, the HMM is too restrictive because: 1) Sampling and training separately for hierarchical events will increase computation complexity; 2) Not being suitable to analyze and recognize high-level event with complicated layers; 3) The interaction across higher semantic levels is not incorporated during modal training; 4) Inability to share structures of lower-level modals. Other extended models of the HMM such as the hierarchical hidden Markov model (HHMM) [2], AHMM[3-4] or statistical date correlation is also used to track trajectory and recognize people's behaviors. All the above approaches cannot respond to environmental change in a real-time and adaptive way.

Moncrieff et al. [5] proposed an anxiety emotion to represent an abnormality in a home environment and implement the abnormality prediction of two interactive devices in a smart home. In fact, the interactions of more than two devices at the same time are more general. We will research interactions of three devices and improve on the probability model to detect abnormalities of multi-devices cross events.

Focusing on addressing issues of environmental context-awareness and real-time abnormality detection in SMS4ENE, we propose a layering representation of architecture as shown in Fig. 1. Various perceptual components capture large amounts of discrete events by recognizing and processing the original video data (such as face recognition, gesture recognition). The HHMM is used to classify these events into different event patterns. And we design an MMO to express knowledge and to share semantics. We establish an improved pessimistic emotion model (PEM) to detect device abnormal behaviors. The system can provide component-based services to front-end users, such as reminder of an occupant, emotion player and customized video search.

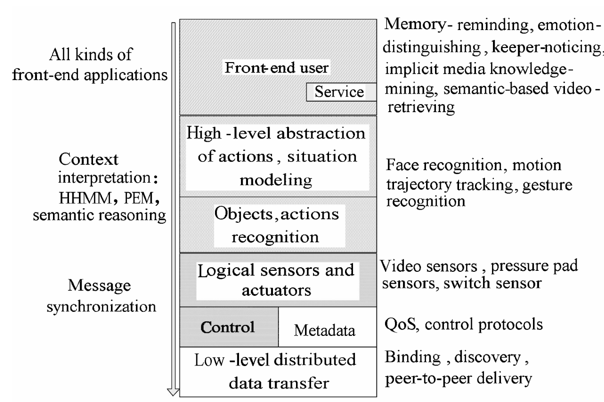


**Fig. 1** Layering representation of architecture

## 1 MMO Model

Corresponding to the layered property of both the HHMM and concepts in real world, we construct a layered MMO by utilizing the method of semantic-layering and semantic-abstracting step by step. It provides a universal specification to annotate and share the semantics for media characters from

**Received** 2008-04-15.
**Biographies:** Yang Lei (1986—), female, graduate; Yang Luming (corresponding author), male, professor, yang@ mail. csu. edu. cn.
**Foundation items:** The National Natural Science Foundation of China (No. 60773110), the Youth Education Fund of Hunan Province (No. 07B014).
**Citation:** Yang Lei, Yang Luming, Man Junfeng, et al. Detecting abnormalities for empty nest elder in smart monitoring system[J]. Journal of Southeast University (English Edition), 2008, 24(3): 347 – 350.

low-level to high-level.

MMO includes core ontology (CO), media ontology (MO) and domain ontology (DO), it has properties as follows: 1) Layered semantics; 2) Interoperability; the OWL document described by MMO can be used as the interface between the media information and the processing program data; 3) Expansibility and universality, it can describe media content with universal standards; 4) Reasoning support.

Layering representation of the MMO is shown in Fig. 2. Various perceptual components take advantage of sensor information to recognize the object and action, then they annotate them into semantic objects and semantic action. First, we abstract actions and objects which are related to media knowledge into sub-events, and then abstract sub-events again into high-level events with HHMM. And finally we annotate the understood content about media with MMO, which is convenient for semantic-based reasoning and retrieval in the future.

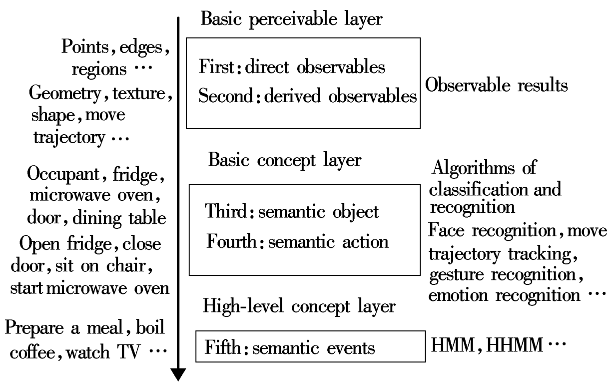


**Fig. 2** Layering representation of MMO

## 2 Abnormal Event Detection

### 2.1 Occupants abnormal behavior detection

The HHMM is a model that integrates the semantics across different layers. It can handle hierarchical modeling of complex dynamic processes of events. So we use it to learn human event patterns. Parameter estimation is implemented with an extended Baul-Twelth algorithm[4].

By learning from a large number of sample sequences which are the results of processed video data collected over a very long period, we select a model of the event "prepare a meal" shown in Fig. 3 to illustrate the application of the HHMM in abnormality detection and the MMO application in annotation. In the HHMM model, "prepare a meal" is a high-level event abstracted by learning with the HHMM, it corresponds to a semantic event (SE). This event is subdivided and trained from sub-events named "wash", "cook" and "bake" which consist of low-level observational activity sequences. These sub-events are annotated with concept semantic sub-actions (SSA). How the MMO is mapped to the HHMM from low-level to high-level is shown in Fig. 3. And the annotation semantic relationships of the high level events are shown in Fig. 4.

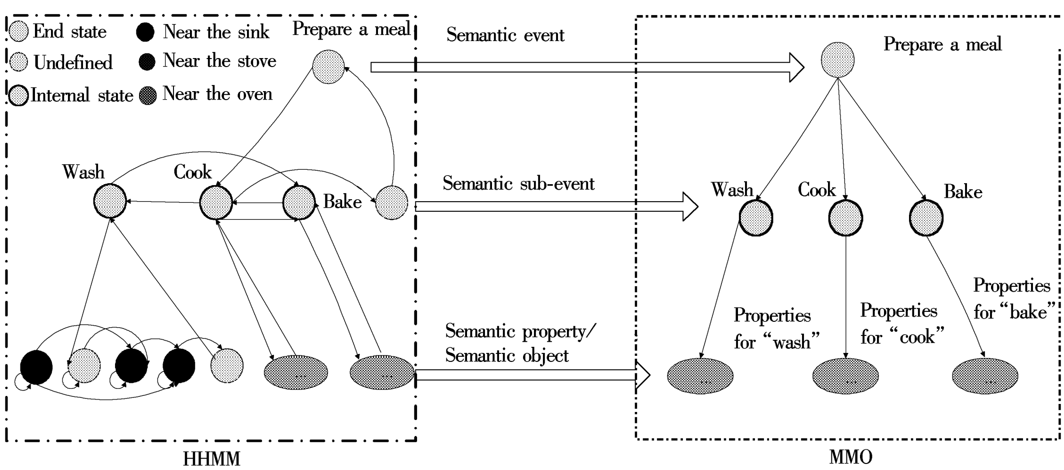


**Fig. 3** The mapping relationship between HHMM and MMO

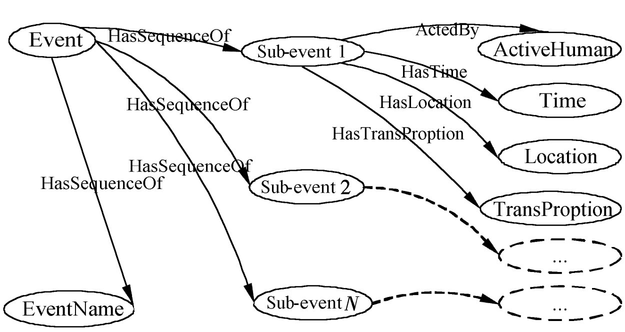


**Fig. 4** Annotation semantic relationship of high-level events

The classifications of normality and abnormality are a problem of HHMM model estimation. The estimation of observation sequences is implemented with the EM algorithm[6]. The computation formula for abnormality classifications is as follows:

$$P(o_1, o_2, \dots, o_t \mid \pi, \lambda) > P_{\text{thred}} \tag{1}$$

The current observation sequences are compared to sequences in all the event-models which have been trained with the HHMM. If the similarity value exceeds a threshold value $P_{\text{thred}}$, we classify the current sequences to the very model that has the similarity value $P_{\text{thred}}$ with the current sequences; if the current observations cannot be classified to any model of the HHMM models, then abnormality is detected.

### 2.2 PEM and device abnormality detection

Although using video real-time surveillance and learning of behavior patterns is feasible, it is difficult to avoid errors due to inaccurate context-awareness about the environment, especially when the object is sheltered or occluded or the action is too light to recognize. So we present a new pessimistic emotion model (PEM) to detect activity abnormalities of multi-devices cross events.

### 2.2.1 The idea of PEM

We hope that SMS4ENE should simulate rational thinking and sensible emotion judgment as humans do to detect device abnormalities. The PEM uses the concept of "pessimistic emotion" to represent the cognition of states of hazardous devices, and the pessimistic emotion value (PEV) to describe the degree of device abnormality. The PEV is zero when a device is turned on, but rises over time if it is not interacted with or if it is ignored by the occupant. Eventually when it reaches a threshold, some action should be taken.

### 2.2.2 The probability model of PEM

The core idea of the PEM model is to use a measurable value PEV to detect whether the active device is abnormal. The time context is the most important clue representing device activity. Thus, we use the agent-based MJS to acquire device activity and record context time, and then analyze statistical probability of multi-device cross events.

We consider an example sequence as follows: 1) Turn on an oven to make a roast; 2) Turn on the stove to burn oil; 3) Walk to the sink to wash vegetables; 4) Walk to stove; fry vegetables; 5) Turn off the stove. In these sequences, stove and oven are hazardous devices if they are ignored too long. The temporal sequence of cross events of multi devices about the PEM is shown in Fig. 5.

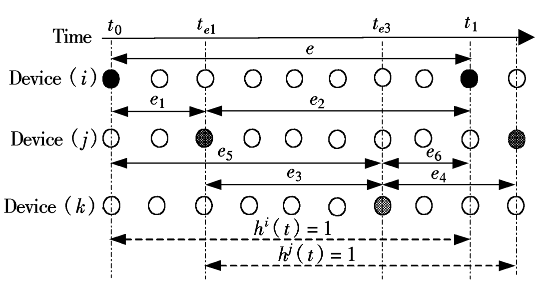


**Fig. 5** Temporal sequence of cross events of multi-devices

$D^i(t_0)=1$ and $D^i(t_1)=1$ indicate an interaction with device $i$ at time $t_0$ and $t_1$, respectively; $h^i(t)=1$ indicates that device $i$ is in a hazardous state. When device $i$ is in a hazardous state, device $j$ is started at $t_{e1}$ and device $k$ at $t_{e3}$. Because passive device $j$ and $k$ are interactive between $t_0$ and $t_1$, the rising speed of the PEV with regard to device $i$ becomes slower as well as the passive device $k$. Here, we define four models to represent PEM.

1) Self interaction duration model: $P^i_{\mathrm{Self}}(e_0)$ denotes the cumulative distribution of the time intervals between interactions with device $i$;

2) Inner interaction duration model: $P^{i,j}_{\mathrm{IID}}(e_2)$ captures the correlation that when passive device $j$ is interactive with. While device $i$ is in a hazardous state, device $i$ will be interacted with again;

3) Inner action duration model: $P^{i,j}_{\mathrm{IAD}}(e_1)$ captures the correlation between a device in a hazardous state and potential interactions with other devices;

4) Interaction event model: $P^{i,j}_{\mathrm{IE}}$ denotes the cumulative distribution of interaction between the user and passive device $j$ while device $i$ is in a hazardous state.

The first, second and third models consider the factor of an occupant's location. If the distance between the occupant and the devices is very close, the value of l is less, otherwise it is greater. The scaling factor associated with each device is decided by the second, third and fourth model:

$$S^{i,j}(t_0,t_{e_j}) = 1.0 - P^{i,j}_{\mathrm{IE}}(1.0 - P^{i,j}_{\mathrm{IAD}}(t_{e_j} - t_0))(1.0 - P^{i,j}_{\mathrm{IID}}(t - t_{e_j})) \tag{2}$$

where $t_{ej}$ is the last time interaction with device $j$.

$$S^{j,k}(t_{el},t_{e_k}) = 1.0 - P^{j,k}_{\mathrm{IE}}(1.0 - P^{j,k}_{\mathrm{IAD}}(t_{e_k} - t_{el})) \cdot (1.0 - P^{j,k}_{\mathrm{IID}}(t - t_{e_k})) \tag{3}$$

$$S^{i,j,k} = r_1 S^{i,j} + r_2 S^{i,k} \tag{4}$$

where $t_{ek}$ is the last time interaction with device $k$, $r_1$ and $r_2$ are effect factors of passive devices $j$ and $k$ to $i$.

$$\mathrm{PE}^{i,j,k}(t) = P^i_{\mathrm{Self}}(t - t_0)\prod_{\forall e_j, e_k} S^{i,j,k}(t_0, e_1, e_2, e_5, e_6) \tag{5}$$

$$\mathrm{PE}^{j,k}(t) = P^j_{\mathrm{Self}}(t - t_{el})\prod_{\forall e_k} S^{j,k}(t_{e1}, e_3, e_4) \tag{6}$$

When three devices interact with each other, we should calculate two PEV. If their values are higher than the threshold, the system will take measures, e. g., the system will inform occupants or keepers by PDA.

## 3 Experimental Results

By training and analyzing a large amount of data collected in the simulation environment, the overall number of the elderly's action events in a day is collected. By using the HHMM to abstract discrete action sequences which is recognized by the video component, we obtain the overall human action amount shown in Fig. 6, and the event amount of devices shown in Fig. 7 is statistical information collected by sensors of MJS.

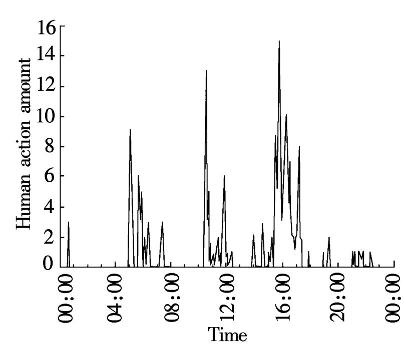


**Fig. 6** The amount of human action events

Fig. 8 shows the system detection error rate. The real line represents the error rate when abnormalities are only detected with HHMM, and the dashed line shows the error rate when the system introduces the improved PEM to enhance the accuracy of abnormality detection for devices. The error rate is obviously lower than the one when it only uses the

HHMM to detect abnormalities.

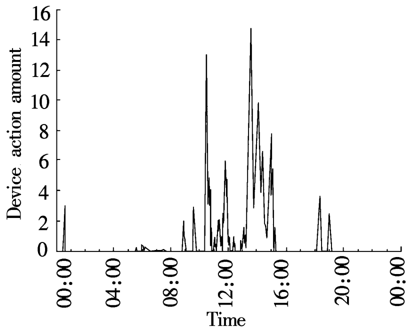


**Fig. 7** The amount of device action events

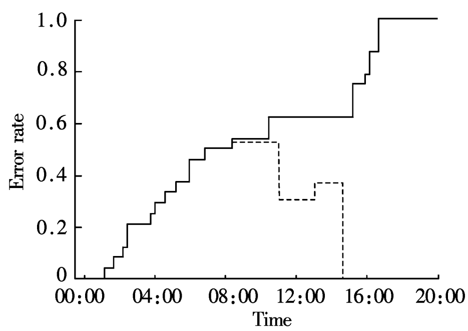


**Fig. 8** Event detection error rates

## 4 Conclusion

This paper discusses how to detect abnormalities of the elder and devices in SMS4ENE, the HHMM is used to learn human behavior patterns and to judge abnormal behavior, and the improved PEM is integrated to implement real-time unsupervised abnormality detection for devices. The MMO is a knowledge expression interface between data of video components and varieties of sensing data; it allows the context knowledge to be shared smoothly and provides reliable and timely clues for reasoning. The MMO model can also be applied in other smart environments.

In future research, we will implement a customized retrieval of media content, and continue to improve the precision of abnormality detection and context-awareness.

## References

[1] Kijak E, Oisel L, Gros P. Hierarchical structure analysis of sport videos using HMMs[C]//*International Conference on Image Processing*. Barcelona, Spain, 2003: 1025 – 1028.

[2] Haussler David. The hierarchical hidden Markov model: analysis and applications [J]. *Machine Learning*, 1998, **32**(1): 41 – 62.

[3] Nguyen N T, Venkatesh S, Bui H. Recognising behaviours of multiple people with hierarchical probabilistic model and statistical data association [C]//*BMVC*2006. Edinburgh, 2006: 1229 – 1239.

[4] Nguyen N T, Phung D Q, Venkatesh S, et al. Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model [C]//*Proc IEEE Conf Computer Vision and Pattern Recognition*. San Diego, CA, USA, 2005: 955 – 960.

[5] Moncrieff Simon, Venkatesh S, West Geoff, et al. Multi-modal emotive computing in a smart house environment [J]. *Pervasive and Mobile Computing*, 2007, **3**(2): 74 – 94.

[6] Jank Wolfgang. The EM algorithm, its stochastic implementation and global optimization: some challenges and opportunities for OR [J]. *Computer Science Interface*, 2006, **36**(3): 367 – 392.

# 空巢老人智能监护系统中异常检测问题

杨 蕾[1] 杨路明[1] 满君丰[1,2] 刘广滨[2]

([1] 中南大学信息科学与工程学院,长沙 410083)
([2] 湖南工业大学计算机与通信学院,株洲 412008)

**摘要:**为实现空巢家庭内老人和家用设备异常行为的实时预测,用多模态传感器获取行为的离散动作序列,并用改进的多层隐马科夫模型抽象出人的高层行为——事件,从大量的时空数据中形成描述居住着正常行为的结构化表达模型,这些模型用作检测居住者异常行为的分类器.为表达推理预测所需的环境上下文信息,设计了多媒体本体(MMO)来标注和推理智能监护系统中的媒体信息.改进了一种悲观情感模型(PEM)来分析室内多活动设备的多交叉事件.实验证明,当被检测的设备处于盲区或被遮挡的情况下,PEM能增强对活动设备检测的准确性和可靠性,上述方法在异常的实时检测方面有很好的性能.

**关键词:**多媒体本体;语义标注;异常检测;分层隐马尔科夫模型;悲观情感模型

**中图分类号:**TP391