

# Mobile humanoid agent with mood awareness for elderly care

Tan, Ah-Hwee; Wang, Di

2014

Wang, D., & Tan, A.-H. (2014). Mobile humanoid agent with mood awareness for elderly care. 2014 International Joint Conference on Neural Networks (IJCNN), 1549-1556.

<https://hdl.handle.net/10356/101364>

<https://doi.org/10.1109/IJCNN.2014.6889916>

---

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [<http://dx.doi.org/10.1109/IJCNN.2014.6889916>].

*Downloaded on 27 Apr 2025 09:49:27 SGT*

# Mobile Humanoid Agent With Mood Awareness for Elderly Care

Di Wang and Ah-Hwee Tan

*Abstract*—Human, especially elderly, require frequent attention, continuous companionship, and deep understanding from the others. To provide more specific and appropriate tender care to the elderly, knowing their affective states is a great advantage. Recent work on human emotion recognition shows promising results that the expressive emotion can be successfully captured through visual, audio, and keyboard or touchpad stroke pattern signals. Furthermore, human activities are shown to be accurately recognizable with context by non-intrusive sensors within or connected to the smartphones. In this paper, we propose a computational model to characterize the affective states of the elderly based on the recognizable daily activities. Therefore, by integrating such an understanding module into a humanoid agent residing in the smartphone platform, we make the mobile agent more human-like. The initial knowledge of the activity-affect associations is taken from published work in psychology and gerontology. Based on the provided training signals, our model adapts the activity-affect knowledge accordingly. Consequently, by modeling mood awareness of the elderly, our agent can carry out more specific task and provide more appropriate tender care.

## I. INTRODUCTION

**E**MOTION recognition has become an emerging research field in recent years. It is an important and challenging inter-disciplinary topic involving psychology and many computerized methods, such as facial and speech recognition, activity recognition with context awareness, and ambient intelligence. To model mood awareness of the elderly, we should also bring gerontology into the overall framework.

The world population is aging fast and it is the elderly who require more attention, companionship, and most importantly understanding from the others. According to United Nation [1], the number of people aged 60 and above is expected to increase from 841 million in 2013 to 2 billion in 2050. We should better prepare now the technologies to assist the elderly. Thus, sensing the affective states of the elderly becomes more important because we can provide better tender care based on proper mood awareness.

Some existing emotion recognition technologies are restricted within certain low-noise confined environments. With recent advances in ambient intelligence, wearable sensors, and mobile communication technologies, the means to model mood awareness are further broadened. In this paper, we propose a Humanoid Agent With Mood Awareness (**HAWMA**) residing in the smartphone platform to provide better and constant tender care to the elderly.

Authors are with School of Computer Engineering, Nanyang Technological University, Singapore (email: {wangdi, asahtan}@ntu.edu.sg).

This research is supported in part by Interactive and Digital Media Programme Office (IDMPO), National Research Foundation (NRF) hosted at Media Development Authority (MDA) under Grant No.: MDA/IDM/2012/8/8-2 VOL 01.

There are many reasons that we select smartphones as the platform to deploy HAWMA. First of all, the current penetration rate of smartphones is high in many developing and developed countries. For example, it is estimated that 83% of Singaporean aged 55 and above own at least one smartphone [2]. Moreover, smartphones nowadays are powerful enough to support natural human-machine interaction and to assemble various kinds of information through an increasing number of build-in and wirelessly connectable sensors. In addition, smartphones provide more convenience: 1) simplicity to monitor the social interactions through text messages, social network service platforms, and other means of communications, if authorized; 2) multimedia platform to support audio, video, and other types of inputs and outputs; 3) an easier way for the elderly to explicitly annotate their affective states through an emoticon system [3].

In this paper, we focus on how to model the affective states of the elderly based on their recognizable activities so that HAWMA can provide more specific and appropriate tender care. Mood awareness takes two major steps: identifying the activity and then determining how it affects the elderly emotionally. Although the data set used in this paper is artificial, we argue that all activities are recognizable by smartphones, computers, and robots with mobility and visual inputs, if authorized. A widely studied and applied emotional model, which maps eight major affective states (evenly distributed) in a 2-D space, is employed to characterize the mood of the elderly. The relationship between the two axes of the emotional model is orthogonal, so they can be independently modeled during computation. If we define the Activity-Affect Knowledge (**AAK**) as the set of all associations between recognizable activities and how they affect the corresponding mood changes in the 2-D space, we can apply such knowledge to model the mood awareness in a computational way. The initial AAK is taken from published work in psychology and gerontology. HAWMA employs a neural network to store, retrieve, and adapt AAK according to the provided training signals with mathematical equations and learning algorithms. After adequate training, HAWMA is able to capture the personalized AAK for different individuals. Consequently, more specific task can be carried out to provide more appropriate tender care.

Although the subjects of this paper are the elderly, the mood awareness model can be generally adopted to other people, even other humanoid agents if all required interfaces are available. Furthermore, to the best of our knowledge, our work is the pioneer to explore how to model mood awareness (specifically) of the elderly through a humanoid agent residing in the smartphone platform.

The rest of the paper is organized as follows. Section II

reviews the related work. Section III introduces the employed emotional model and the pre-defined activity-affect knowledge. Section IV presents the employed network to store, retrieve, and adapt the activity-affect knowledge. Section V shows the experimental results with visualizations. Section VI concludes this paper and proposes future work.

## II. RELATED WORK

There is a common stereotype that the elderly are not receptive to the advanced technological gadgets. However, Czaja [4] shows that although less confident than younger people, the elderly who have successful experience with modern gadgets generally hold more positive attitude and greater confidence. Furthermore, if computers (extensible to smartphones) are modified according to their specific needs with simpler interfaces that follow the natural mental concepts, the elderly can benefit much from them [4]. Therefore, deploying humanoid agents in smartphones is a practical and mobile way to assist the daily lives of the elderly.

To model the affective states of the elderly, the agent must be aware of their current activities with as much context as possible. A thorough investigation on activity recognition using the embedded sensors within the smartphones is presented in [5], wherein similar models are reviewed and their capabilities are compared. Context awareness can be studied in either the social aspect [6] or the physical environmental aspect [7]. Human behavior cognition [8] generates high-level context based on the recognized low-level activities. In this paper, we model the mood awareness of the elderly by assuming the activities have been accurately recognized.

A body of research work focuses on applying established emotional models to regulate the behaviors or personalities of the agents without sensing the affective states of the subjects to be served. Wu and Miao [9] integrate the curiosity model [10] into the virtual peer-learners to accompany secondary school students in the virtual learning world. Ailiya *et al.* [11] integrate the well-known OCC model [12] into the teaching companion agents whose avatars are rendered as dinosaurs in the virtual learning world. In this paper, we do not investigate how to construct the emotional model of our agent. Instead, we focus on how to model the affective states of the elderly. Because people prefer computers that match their own personalities [13], we can simply let our agent follow the affective states of the elderly.

The two major means to recognize the emotion of the user in real time are by analyzing either speech or visual signals. Attabi and Dumouchel [14] construct various combinations of models (various back-end systems with various front-end systems and various front-end systems only) to recognize one of the five basic emotions from speech. Majumder *et al.* [15] propose an extended Kohonen SOM to identify one of the six basic emotions from 26 dimensional geometric facial feature vectors. In this paper, we do not require the mood of the elderly to be explicitly expressive for sensing.

Smartphone itself can either recognize the emotions of the users or express its own emotions. Lee *et al.* [16] propose a way to identify one of the seven basic emotions of the

users by analyzing both their patterns (speed of the key strokes and counts) when they write Twitter messages and the environmental information (location, time, and weather). Based on the interactions between the user and the smartphone, Kifor *et al.* [17] apply the pre-defined activity-affect knowledge to model the emotion (among five basic ones) of the smartphone. In this paper, we focus on modeling the affective states of the elderly based on their high-level activities recognized. Furthermore, the pre-defined activity-affect knowledge evolves if training signals are provided.

In recent years, research work on caring the elderly with emotional supports is rising in numbers. McCalley and Mertens [18] develop a pet plant, which has no hygienic concerns, to attract the elderly's attention and to follow their extrovert or introvert characters. Although the evaluation results are significant in terms of therapeutic effects, the pet plant is still limited in functionality. It cannot move or express itself other than by means of flashing lights in different colors and manners. On the other hand, an agent residing in the smartphone platform is mobile and can naturally interact with the elderly through various multimedia applications. Zhou *et al.* [19] propose a promising elderly-orientated smart-home environment. However, within their framework, there is no agent implemented to provide human-like personal assistance. In this paper, we show how our agent provides better tender care when the elderly is experiencing an emotional extreme.

## III. COMPUTATIONAL ACTIVITY-AFFECT KNOWLEDGE

Elderly tend to respond more to the positive stimuli than the negative ones [20]. Generally speaking, elderly require constant positive companionship to keep them away from loneliness, depression, and anxiety. Because different activities performed or experienced by the elderly affect their mood in different ways, the Activity-Affect Knowledge (AAK) must be carefully modeled for accurate computation (see Section III.A) and efficient storage, retrieval, and adaptation (see Section IV). Furthermore, the AAK is different among different elderly [21], evolves as they age [22], and changes when they have different roles [23]. Therefore, the AAK should be presented with a simple yet effective model to characterize different temperaments. In this paper, our proposed agent, named Humanoid Agent With Mood Awareness (HAWMA), employs the Russell's circumplex model of affect [24], which is one of the most widely studied and applied emotional models [25].

### A. Representation of the affective states

Russell finds that the relationship between pleasure-misery and arousal-sleepiness is almost orthogonal [24]. Therefore, he proposes a circumplex model (see Fig. 1) to represent one's affective states in the 2-D space. The eight major ones (represented in bold) are evenly distributed around the circle. By knowing the exact pleasure and arousal values, one of the eight affective states can be identified accordingly.

The Pleasure-Arousal-Dominance (**PAD**) model [26] extends the 2-D model [24] by affixing another axis to represent

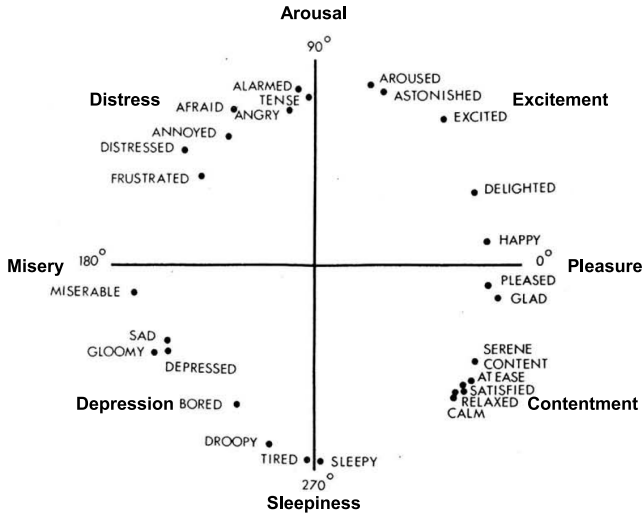


Fig. 1. Russell's circumplex model of affect.

dominance. The positive or negative values of dominance represent how much a person is affected by the activity, in dominance or in submissiveness, respectively.

HAWMA uses the dominance values to determine mathematically how long the activities affect the mood of the elderly. Because each individual activity may affect the pleasure and arousal values differently, we separate the two axes for independent computation. Furthermore, because the influence of previous activities decreases as time elapses, we employ the one-tailed Gaussian function to mathematically model how each activity affects the mood:

$$f(x) = -a \exp\left(-\frac{(t - t^x)^2}{2b^2}\right), t \geq t^x, \quad (1)$$

where  $f(x)$  denotes the function to compute the amount of affect contributed by the activity happened at time  $t^x$  in either the pleasure-misery or the arousal-sleepiness dimension at time  $t$ ;  $a$  denotes the altitude of activity  $x$  in terms of affect; and  $b$  denotes the corresponding effective time (bandwidth) of activity  $x$ . As  $t \geq t^x$ , the amount of affect contributed by activity  $x$  decreases as  $t$  increases.

Eq. (1) models how much each individual activity happened preceding the current time  $t$  contributes to the current pleasure or arousal values. At any time  $t$ , the actual pleasure or arousal value is defined as the aggregation of all values contributed by the previously happened activities:

$$P_t = \sum f_P(x), \forall t^x \leq t; A_t = \sum f_A(x), \forall t^x \leq t. \quad (2)$$

After computing  $P_t$  and  $A_t$ , HAWMA can sense the affective state of the elderly in the 2-D space shown in Fig. 1 in real time. If we define the circle centered at the origin with a radius of 1 to represent the safety boundary, then HAWMA can notify any concerned party (including the elderly) about the sensed extreme emotional situation (the affective state falls outside the safety boundary). This functionality is defined in Eq. (3).

$$\text{IF } \sqrt{(P_t^2 + A_t^2)} > 1, \\ \text{THEN send corresponding notification(s).} \quad (3)$$

TABLE I  
EXAMPLES OF THE PRE-DEFINED ACTIVITY-AFFECT KNOWLEDGE.

ID	Activity	PAD Values
A.71	Hear that an old friend got cancer.	-0.15; 0.58; -1.44
B.1	Wake up from a good sleep.	0.75; 1.00; 1.50
C.13	Keyword "miserable" is identified.	-0.99; -0.15; N/A

### B. Pre-defined activity-affect knowledge

The pre-defined set of Activity-Affect Knowledge (AAK) used in this paper is taken or derived from the published work in psychology and gerontology. There are three categories of the pre-defined AAK: A) averaged survey data consisting of PAD values with the corresponding activities; B) heuristic rules derived from expert knowledge; and C) training signals associated with keywords. One example from each AAK category is presented in Table I.

Knowledge in Category A is imported from [27], which reports the averaged survey data collected from 200 young adults on emotional correlates. In [27], there are 77 unique activity-affect pairs reported. Among all 77 pairs, some activities cannot be recognized based on the currently available sensors. However, 20 activities can be recognized by smartphones, computers, and robots with mobility and visual inputs, if authorized. Moreover, 14 activities can be recognized if they are approximated by closely similar activities. Because the survey uses five-point Likert scale, each PAD value is bounded within the  $[-2, 2]$  interval. Rule A.71, whose wording is shortened in Table I, can be directly recognized if someone tells the elderly this message through Twitter. Because the survey data is not collected from the elderly, the misery value ( $P$ ) seems to be low. However, we show in Section V that the pre-defined knowledge evolves after adaptations if training signals are provided.

Knowledge in Category B consists of heuristic rules derived from expert knowledge. McCrae *et al.* [28] show that the positive or negative affective state highly correlates with how well the elderly scale last night's sleep. In [28], Positive Affect and Negative Affect Scales (PANAS) [29] are used. However, PANAS values can be translated [30] to PAD values. The positive correlate is illustrated as Rule B.1 in Table I that a good sleep suggests a high pleasure value, a higher arousal value, and both of them last for a relatively long period of time. The PAD values in Category B have the same scale ( $[-2, 2]$ ) as those in Category A.

Knowledge in Category C consists of all the training signals when certain keyword is identified. The keywords can be identified through parsing all outgoing messages and explicitly provided by the elderly through the annotating emoticon functionality offered by HAWMA residing in the smartphone platform. Among all the 28 keywords shown in Fig. 1, 15 of them are given with the exact degrees in [24]. Assuming the length of the segment is 1 (the scale boundary

is  $[-2, 2]$ ), the exact  $P$  and  $A$  values ( $\cos(\theta)$  and  $\sin(\theta)$ ) can be computed as the training signals. Because the values are used for training only, the  $D$  values are not applicable here.

All the pre-defined AAK is listed in details (how to recognize the activities, how activities are approximated by closely similar ones, how heuristic rules are derived, and how to estimate the values of the training signals if the exact degrees are not provided) in an Excel file<sup>1</sup>. If we denote these pre-defined values as raw values, the parameters (altitude  $a$  and bandwidth  $b$ ) of Eq. (1) are defined as follows.

$$P_a = P_{raw}/2; A_a = A_{raw}/2. \quad (4)$$

$$P_b = A_b = 8|D_{raw}|, \quad (5)$$

where  $|y|$  denotes the absolute value of  $y$ ; and the coefficient is set to 8 so that the maximal duration of any activity approximately affects the elderly for 4 hours (time is discretized into 15-minute intervals).

Although the initial values of  $P_b$  and  $A_b$  are identical, we show in Section V that they are differentiable after training.

#### IV. ACTIVITY-AFFECT MODEL FOR KNOWLEDGE STORAGE, RETRIEVAL, AND ADAPTATION

Section III shows how to use AAK to compute the  $P$  and  $A$  values in real time. In this section, we propose a neural network, named Activity-Affect Model (**AAM**), to store, retrieve, and adapt AAK for each individual elderly.

##### A. Dynamics of the activity-affect model

The architecture of AAM is shown in Fig. 2. AAM is a particular realization of the generic fusion ART network [32], which applies the Adaptive Resonance Theory (ART) [33] for learning. AAM employs a 3-channel architecture, comprising a high-level association field wherein the associations among the three low-level fields are stored as cognitive codes. All the pre-defined AAK is translated (see Eqs. (10) and (11)) and inserted into AAM during initialization. AAM supports bottom-up knowledge storage and top-down knowledge retrieval. An important consideration that we employ the fusion ART network to model mood awareness is because its network structure can dynamically evolve by either adding in new cognitive codes in  $F_2^c$  or expanding the length of the input vectors of the three low-level fields after initialization. More details about another realization of the fusion ART network can be found in [34]. In this paper, AAM employs ART2 operations [35], [36] for knowledge adaptations. The dynamics of AAM are described as follows.

**1) Input vectors:** The input vector  $I^{c1}$  to  $F_1^{c1}$  consists of six binary bits:  $\{I_i^{c1} \text{ for } i = 1, \dots, 6\}$ . Therefore,  $I^{c1}$  indicates the unique index of each activity (totally 39 defined). As discussed earlier in Section III.A, the affect knowledge is splitted into the independent pleasure ( $P$ ) and arousal

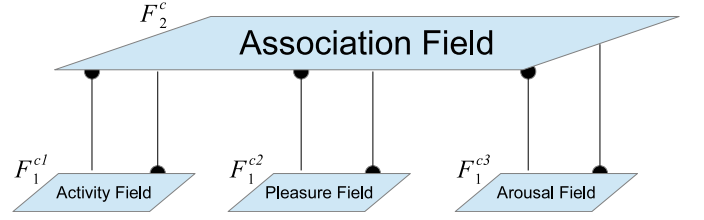


Fig. 2. Network architecture of the Activity-Affect Model (AAM).

( $A$ ) values. Therefore, both  $I^{c2}$  and  $I^{c3}$  consist of two real numbers representing the altitude  $a$  and bandwidth  $b$  parameters defined in Eq. (1), respectively.

**2) Activity vectors:** Let  $\mathbf{x}^{ck}$  denote the  $F_1^{ck}$  activity vector for  $k = 1, 2, 3$ . Let  $\mathbf{y}^c$  denote the  $F_2^c$  activity vector.

**3) Weight vectors:** Let  $\mathbf{w}_j^{ck}$  denote the weight vector associated with the  $j$ th code in  $F_2^c$  for learning the input patterns in  $F_1^{ck}$  for  $k = 1, 2, 3$ .

**4) Parameters:** The dynamics of AAM is determined by learning rate parameters  $\beta^{ck} \in [0, 1]$  for  $k = 1, 2, 3$ ; contribution parameters  $\gamma^{ck} \in [0, 1]$  for  $k = 1, 2, 3$ , where  $\sum_{k=1}^3 \gamma^{ck} = 1$ ; and vigilance parameters  $\rho^{ck} \in [0, 1]$  for  $k = 1, 2, 3$ .

**5) Code activation:** By representing the low-level input vectors for bottom-up retrieval, the activation value of each  $F_2^c$  code  $j$  is computed by the choice function  $T_j^c$ :

$$T_j^c = \sum_{k=1}^3 \gamma^{ck} \frac{\mathbf{x}^{ck} \cdot \mathbf{w}_j^{ck}}{\|\mathbf{x}^{ck}\| \|\mathbf{w}_j^{ck}\|}, \quad (6)$$

where the operation  $\cdot$  is the dot product and the norm  $\|\cdot\|$  is defined by  $\|\mathbf{p}\| \equiv \sqrt{\sum_i p_i^2}$ .

**6) Code competition:** After the activation values of all codes in  $F_2^c$  are computed with respect to the current input vectors, the winner code  $J$  in  $F_2^c$  is selected as the one who has the maximal activation value:

$$T_J^c = \max\{T_j^c : \text{for all } F_2^c \text{ code } j\}. \quad (7)$$

For knowledge retrieval, the process ends after the values stored in  $w_j^{ck}$  are read out. For knowledge adaptation, further computation is required.

**7) Template matching:** Before code  $J$  can be used for learning, a template matching process checks that  $w_j^{ck}$  are sufficiently close to  $x^{ck}$  for  $k = 1, 2, 3$ . Specifically, resonance occurs if for each channel  $k$ , the match function  $m_J^{ck}$  of the chosen code  $J$  meets its vigilance criterion:

$$m_J^{ck} = \frac{|\mathbf{x}^{ck} \wedge \mathbf{w}_J^{ck}|}{|\mathbf{x}^{ck}|} \geq \rho^{ck}. \quad (8)$$

**8) Template learning:** Once code  $J$  is selected, for each channel  $k$ , the weight vector  $\mathbf{w}_J^{ck}$  is updated towards the input vector  $\mathbf{x}^{ck}$ :

$$\mathbf{w}_J^{ck(\text{new})} = (1 - \beta^{ck})\mathbf{w}_J^{ck(\text{old})} + \beta^{ck}\mathbf{x}^{ck}. \quad (9)$$

Because the fusion ART network [32] requires all vectors to be bounded within the  $[0, 1]$  interval, the transformations

<sup>1</sup>All the pre-defined activity-affect knowledge and all the generated artificial daily lives are uploaded to our research center's repository: <http://www.ntulily.org/>. The online user evaluations of the generated artificial daily lives will be available soon (currently under construction).

TABLE II  
THE DETAILED ALGORITHM FOR THE OVERALL KNOWLEDGE ADAPTATION PROCESS.

- 
- Step 1 At each time  $t$  when training signals are provided, check  $P_t$  &  $A_t$  for whether adaptations are required.
- Step 2 If the degree defined by  $P_t$  &  $A_t$  is deviated over  $\pm 22.5^\circ$  from the training signal, then rule adaptations are performed in two concurrent criteria: dominance values and recency effects.
- Step 3 If the degree of the training signal is given as  $m$ , then the amount of change required to learn in  $P$  &  $A$  axes are defined as  $\Delta P_t = \sqrt{(P_t^2 + A_t^2)} \cos(m) - P_t$ ;  $\Delta A_t = \sqrt{(P_t^2 + A_t^2)} \sin(m) - A_t$ .
- Adapt the altitudes of rules according to the dominance values:**
- Step 4 For all previously fired rules, compute their values at time  $t$  for  $P$  &  $A$ :  
 $P_{values,t} = \cup f_P(x), \forall t^x \leq t$ ;  $A_{values,t} = \cup f_A(x), \forall t^x \leq t$ .
- Step 5 Based on the absolute values of  $P_{values,t}$  &  $A_{values,t}$ , select at most seven of them (according to the magic number  $7 \pm 2$  [31]) to identify the sets of rules to be adapted:  $Rule\_ID_{Dominance\_P}$  &  $Rule\_ID_{Dominance\_A}$ .
- Step 6 For all rules selected for adaptation, compute their recency effects:  
 $Recency_{P,x} = t - t^x, \forall x \in Rule\_ID_{Dominance\_P}$ ;  $Recency_{A,x} = t - t^x, \forall x \in Rule\_ID_{Dominance\_A}$ .
- Step 7 Compute the normalized learning values for all rules:  $\Delta P_{t,x} = \frac{1}{|Rule\_ID_{Dominance\_P}|} \left(1 - \frac{Recency_{P,x}}{\sum Recency_{P,\forall x}}\right) \Delta P_t$ ;  
 $\Delta A_{t,x} = \frac{1}{|Rule\_ID_{Dominance\_A}|} \left(1 - \frac{Recency_{A,x}}{\sum Recency_{A,\forall x}}\right) \Delta A_t$ , where  $|S|$  denotes the cardinality of the set  $S$ .
- Adapt the bandwidths of rules according to the recency effects:**
- Step 8 Identify at most seven [31] recently fired sets of rules:  $Rule\_ID_{Recency\_P}$  &  $Rule\_ID_{Recency\_A}$ .
- Step 9 Compute their recency effects:  
 $Recency_{P,y} = t - t^y, \forall y \in Rule\_ID_{Recency\_P}$ ;  $Recency_{A,y} = t - t^y, \forall y \in Rule\_ID_{Recency\_A}$ .
- Step 10 Compute the normalized learning values for all rules:  $\Delta P_{t,y} = \frac{1}{|Rule\_ID_{Recency\_P}|} \left(1 - \frac{Recency_{P,y}}{\sum Recency_{P,\forall y}}\right) \Delta P_t$ ;  
 $\Delta A_{t,y} = \frac{1}{|Rule\_ID_{Recency\_A}|} \left(1 - \frac{Recency_{A,y}}{\sum Recency_{A,\forall y}}\right) \Delta A_t$ , where  $|S|$  denotes the cardinality of the set  $S$ .
- Step 11  $\forall \Delta P_{t,y}$ , if  $P_{t,y}$  has different polarity with  $\cos(m)$ , then  $\Delta P_{t,y} = -|\Delta P_{t,y}|$ , otherwise,  $\Delta P_{t,y} = |\Delta P_{t,y}|$ ;  
 $\forall \Delta A_{t,y}$ , if  $A_{t,y}$  has different polarity with  $\sin(m)$ , then  $\Delta A_{t,y} = -|\Delta A_{t,y}|$ , otherwise,  $\Delta A_{t,y} = |\Delta A_{t,y}|$ .
- Formulate the value vectors for learning:**
- Step 12  $P_{t,x} = P_{t,x} + \Delta P_{t,x}, \forall x \in Rule\_ID_{Dominance\_P}$ ;  $P_{t,y} = P_{t,y} + \Delta P_{t,y}, \forall y \in Rule\_ID_{Recency\_P}$ ;  
 $A_{t,x} = A_{t,x} + \Delta A_{t,x}, \forall x \in Rule\_ID_{Dominance\_A}$ ;  $A_{t,y} = A_{t,y} + \Delta A_{t,y}, \forall y \in Rule\_ID_{Recency\_A}$ .
- Step 13 Apply the binary index of each activity with the above computed vectors to AAM for knowledge adaptations.
- Repeat the learning process recursively until no more learning is triggered.**
- 

TABLE III  
TRANSFORMED VALUES OF THE EXAMPLES SHOWN IN TABLE I.

ID	ART- $P_a$	ART- $P_b$	ART- $A_a$	ART- $A_b$
A.71	0.4625	0.72	0.645	0.72
B.1	0.6875	0.75	0.75	0.75

of PAD values (Eqs. (4) and (5)) are defined in Eqs. (10) and (11), respectively. The transformed values of the examples given in Table I are shown in Table III.

$$ART-P_a = (P_a + 1)/2; ART-A_a = (A_a + 1)/2; \quad (10)$$

$$ART-P_b = ART-A_b = |D_{raw}|/2. \quad (11)$$

### B. Learning algorithm for knowledge adaptation

If the current affective state determined by  $P_t$  and  $A_t$  is out of phase with the provided training signal (falls outside

the  $45^\circ$  sector), the amount of affects contributed by the previously recognized activities should be adapted so that  $P_t$  and  $A_t$  are moving towards the provided training signal. As discussed earlier in Section III.A,  $P$  and  $A$  values are computed independently. Moreover, the altitude parameter  $a$  and the bandwidth parameter  $b$  defined in Eq. (1) should be respectively adapted according the dominance values and the recency effects. Furthermore, not all the previously recognized activities should be adapted. Only a number of them, which contribute the most to  $P_t$  and  $A_t$ , are selected. The chosen number is 7 according to the magic number  $7 \pm 2$  [31], which is the limit of the number of objects or events that human can hold in the working memory.

Each step of the overall learning algorithm are introduced with mathematical equations in Table II. The training signals (see Section III.B, Category C) provided (Step 1) should not be directly transformed and applied to AAM for knowledge adaptations. Instead, the relative amount of the difference (Step 3) between the given training signal and the current

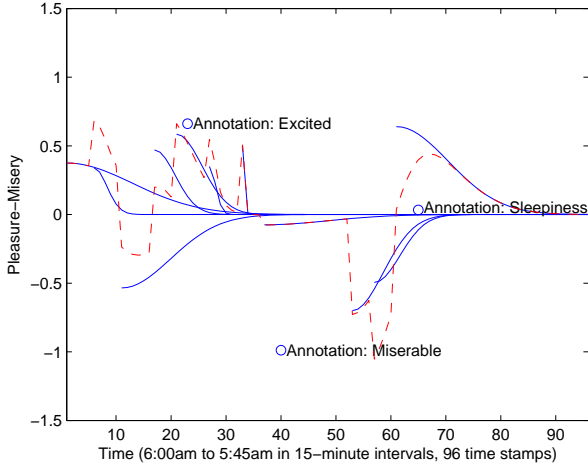


Fig. 3. Affective values in the pleasure-misery dimension before training according to the artificially recognized activities of Elderly-1 in Day-1. The dashed line represents the actual pleasure value  $P_t$ . Each annotation denotes the position of the corresponding training signal. If at time  $t'$ , an activity  $x$  is recognized, then the amount of influence of  $x$  is at the peak (determined by  $a$  defined in Eq. (1)) in time  $t'$ . As time elapses, the influence of  $x$  decreases (the effective duration is determined by  $b$  defined in Eq. (1)). At any time  $t$ ,  $P_t$  (defined in Eq. (2)) is the sum of the individual  $P$  values contributed by all the previously recognized activities.

$P_t$  and  $A_t$  values should be adapted to the selected activity-affect association rules. There are two concurrent selection criteria (Step 2): dominance values and recency effects. The altitude parameters of the rules selected according to the dominance values (Steps 4 to 7) and the bandwidth parameters of the rules selected according to the recency effects (Steps 8 to 11) are adapted in AAM (Steps 12 and 13) according to the respective equations. The learning algorithm runs recursively until no more learning is triggered.

## V. EXPERIMENTAL RESULTS AND VISUALIZATIONS

To show AAM is capable of adapting the pre-inserted AAK, we generate two days of artificial daily lives for two artificial elderly, respectively. Altogether there are 51 activities (with overlaps, taken from either Category A or Category B introduced in Section III.B) and 14 training signals (with overlaps, taken from Category C introduced in Section III.B). The  $P$  and  $A$  values of Elderly-1 in Day-1 are plotted in Figs. 3 and 4, respectively. The corresponding mood transitions of Elderly-1 throughout Day-1 are plotted in Fig. 5.

To maintain the specific AAK according to different individuals, two AAM networks are implemented for the two elderly, respectively. However, the two AAM networks use the same set of parameter values:

- Vigilance parameter  $\rho = \{1, 0, 0\}$ . Because  $F_1^{c_1}$  represents the index of each activity, an exact match is required (i.e.  $\rho^1 = 1$ ). Because each activity only associates with one set of  $P$  and  $A$  values,  $\rho^2 = \rho^3 = 0$ .
- Learning rate parameter  $\beta = \{0, 0.8, 0.8\}$ . Because an exact match is required in  $F_1^{c_1}$ ,  $\beta^1$  can be set to any

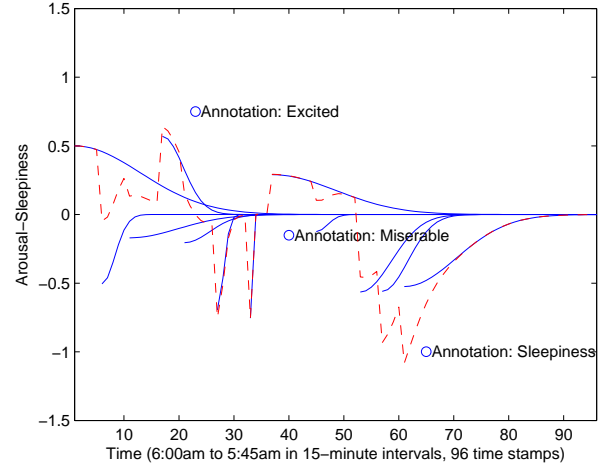


Fig. 4. Affective values in the arousal-sleepiness dimension before training according to the artificially recognized activities of Elderly-1 in Day-1. The dashed line represents  $A_t$  defined in Eq. (2).

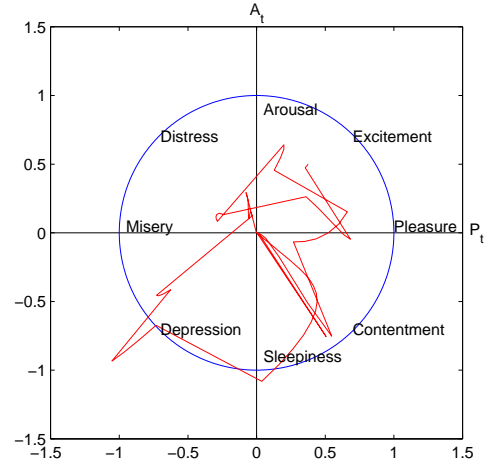


Fig. 5. The mood transitions of Elderly-1 in Day-1 before training. The radius of the circle is 1, which defines the emotional safety boundary.

value in the  $[0, 1]$  interval. Because this experiment is conducted in a relatively stable and recursive manner,  $\beta^2$  and  $\beta^3$  are set to relatively high values for fast learning.

- Contribution factor  $\gamma = \{0.8, 0.1, 0.1\}$ . Because for most of the time, the knowledge retrieval process only depends on the index of the activity,  $\gamma^1$  is much favored.

After applying the learning algorithm shown in Table II, the  $P$  and  $A$  values of Elderly-1 in Day-1 are now plotted in Figs. 6 and 7, respectively. The corresponding mood transitions are plotted in Fig. 8. Moreover, the example rules previously given in Table III are now shown in Table IV.

It is noticeable by comparing Table IV to Table III that Rule A.71 (triggered at time stamp 37 to Elderly-1 in Day-1) is drastically changed. It is shown in Table I that Elderly-1 received a message from Twitter that an old friend got cancer. It is discussed earlier in Section III.B that this kind

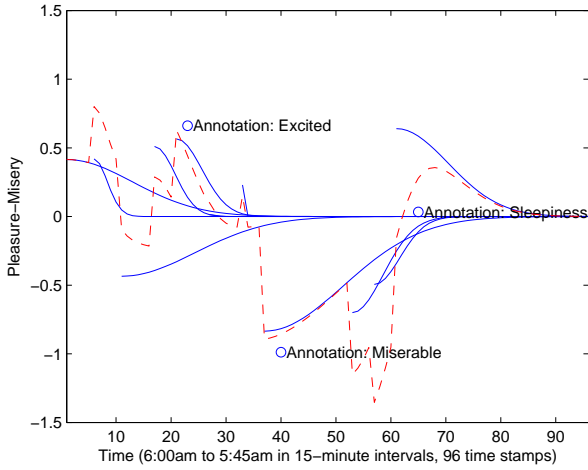


Fig. 6. Affective values in the pleasure-misery dimension after training according to the artificially recognized activities of Elderly-1 in Day-1.

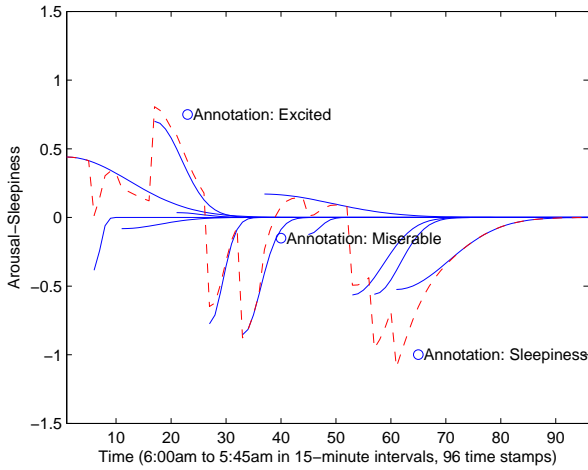


Fig. 7. Affective values in the arousal-sleepiness dimension after training according to the artificially recognized activities of Elderly-1 in Day-1.

TABLE IV  
ADAPTED VALUES OF THE EXAMPLES SHOWN IN TABLE III.

ID	$ART-P_a$	$ART-P_b$	$ART-A_a$	$ART-A_b$
A.71	0.0883	0.862	0.5291	0.747
B.1	0.7042	0.776	0.7692	0.768

of negative sentiment is expected to affect the elderly greatly. Because Elderly-1 posts online: “I just heard... I feel so miserable right now...” at time stamp 40 (45 minutes later), this particular training signal boost up the  $P$  value towards the negative direction.

Comparing Table IV to Table III, Rule B.1 (triggered at time stamp 1) is not significantly changed. However, it is clearly shown that the bandwidth parameters for  $P$  and  $A$  are differentiable despite of the slight changes.

Comparing Fig. 8 to Fig. 5, the line segments representing the mood transitions of Elderly-1 in Day-1 are noticeably

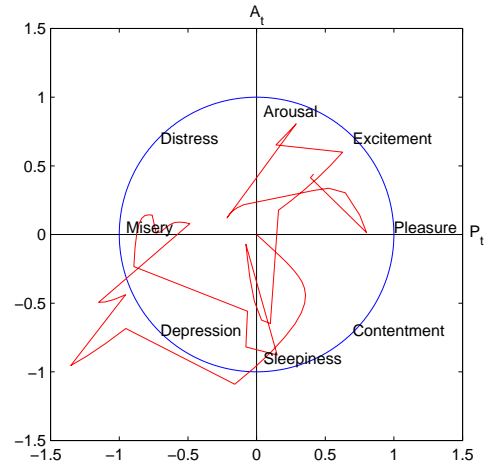


Fig. 8. The mood transitions of Elderly-1 in Day-1 after training.

smoother because there are less sharp and backwards turns.

At any time, whenever the affective state of the elderly moves beyond the circle representing the safety boundary in the emotional aspect, the agent generates a corresponding notification to the elderly. In this context, according to Fig. 5, before knowledge adaptation is performed, the agent sends two different notifications to Elderly-1 in Day-1. According to Fig. 8, after training, the agent sends three different notifications to Elderly-1. Moreover, because multiple warnings are sensed in sequence, the agent also sends a text message to the concerned party, say the daughter of Elderly-1, “Your mother feels depressed earlier around 8:30 pm, would you like to call her or pay her a visit to see whether she is OK now?” Because the affective states of the elderly can be better sensed after training (less sharp or backwards turns and parameter values converges to ensure  $P_t$  and  $A_t$  are closer to and in phase with the training signals), the agent is able to provide more specific and appropriate tender care to the elderly, other than the relatively simple notification services. For example, whenever the agent senses that the elderly is feeling too excited (without exceeding the safety boundary), it may play a peaceful piece of music. Another example is that whenever the elderly feels sleepy (not in the evening around bed time), the agent may suggest activities such as an interesting interactive game to provide some mental stimulation or taking a nap if there is no scheduled event.

## VI. CONCLUSION AND FUTURE WORK

This paper shows how our proposed humanoid agent residing in the smartphone platform models mood awareness of the elderly to provide better tender care. To characterize the affective states in a computational way, the mobile agent employs a neural network to model the activity-affect knowledge for each individual elderly. The initial activity-affect knowledge is taken from published work in psychology and gerontology. However, user-specific knowledge can be captured if adequate training signals are provided. The



transitions of the affective states can be represented in the 2-D space for visualizations. Although the data set in use is artificial, we argue that all the involved activities are recognizable by smartphones, computers, and robots with mobility and visual inputs, if authorized. Furthermore, we show that by employing the computational model for real-time mood awareness of the elderly, the agent is capable of providing more specific and appropriate tender care to support the elderly in the emotional aspect.

This paper is the first step towards the ultimate goal of actual deployments. Numerous challenges are waiting ahead along the rough road. There are mainly two directions to approach the ultimate goal: performing more reliable activity recognition with richer context (including real-time health monitoring through designated wearable gadgets with wireless connectivity) and introducing more capable and dynamic models to deal with the naturally complex and fast-changing human minds.

Other plausible future work includes taking the mental illness history (if any) and the current mental health status of the elderly into consideration, collecting or generating more daily lives of the elderly for a longer duration to illustrate the differences among them, and deploying the agent in a virtual world or a game environment for case studies.

#### REFERENCES

- [1] "World population prospects," United Nation, accessible till January 2014. [Online]. Available: <http://www.un.org/>
- [2] "Smartphones in Singapore: A whitepaper release," Blackbox Research, accessible till January 2014. [Online]. Available: <http://www.blackbox.com.sg/>
- [3] E. Ito and T. Fujimoto, "A proposal of intuitive and immediate emoticons system to do non-verbal communication with smartphones," in *Proceedings of International Conference on Intelligent Systems, Modelling and Simulation*, 2013, pp. 335–339.
- [4] S. J. Czaja, "Computer technology and the older adult," in *Handbook of Human-Computer Interaction (Second Edition)*, M. G. Helander, T. K. Landauer, and P. V. Prabhu, Eds. Elsevier, 1997, pp. 797–812.
- [5] H. Martin, A. M. Bernardos, J. Iglesias, and J. R. Casar, "Activity logging using lightweight classification techniques in mobile devices," *Personal and Ubiquitous Computing*, vol. 17, no. 4, pp. 675–695, 2013.
- [6] E. Miluzzo, C. T. Cornelius, A. Ramaswamy, T. Choudhury, Z. Liu, and A. T. Campbell, "Darwin phones: the evolution of sensing and inference on mobile phones," in *Proceedings of International Conference on Mobile Systems, Applications, and Services*, 2010, pp. 5–20.
- [7] P. Korpipaa, J. Mantyjari, J. Kela, H. Keranen, and E.-J. Malm, "Managing context information in mobile devices," *Pervasive Computing*, vol. 2, no. 3, pp. 42–51, 2003.
- [8] L. Pei, R. Guinness, R. Chen, J. Liu, H. Kuusniemi, Y. Chen, L. Chen, and J. Kaistinen, "Human behavior cognition using smartphone sensors," *Sensors*, vol. 13, no. 2, pp. 1402–1424, 2013.
- [9] Q. Wu and C. Miao, "Modeling curiosity-related emotions for virtual peer learners," *IEEE Computational Intelligence Magazine*, vol. 8, no. 2, pp. 50–62, 2013.
- [10] G. Loewenstein, "The psychology of curiosity: A review and reinterpretation," *Psychological Bulletin*, vol. 116, no. 1, pp. 75–98, 1994.
- [11] Ailiya, Z. Shen, and C. Miao, "An emotional agent in virtual learning environment," *Transactions on Edutainment IV, Lecture Notes in Computer Science*, vol. 6250, pp. 22–33, 2010.
- [12] A. Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.
- [13] T. W. Bickmore and R. W. Picard, "Establishing and maintaining long-term human-computer relationships," *ACM Transactions on Computer-Human Interaction*, vol. 12, no. 2, pp. 293–327, 2005.
- [14] Y. Attabi and P. Dumouchel, "Anchor models for emotion recognition from speech," *IEEE Transactions on Affective Computing*, vol. 4, no. 3, pp. 280–290, 2013.
- [15] A. Majumder, L. Behera, and V. K. Subramanian, "Emotion recognition from geometric facial features using self-organizing map," *Pattern Recognition*, vol. 47, no. 3, pp. 1282–1293, 2014.
- [16] H. Lee, Y. S. Choi, S. Lee, and I. P. Park, "Towards unobtrusive emotion recognition for affective social communication," in *Proceeding of IEEE Consumer Communications and Networking Conference*, 2012, pp. 260–264.
- [17] T. Kifor, T. Gottdank, A. Hajnal, P. Baranyi, B. Korondi, and P. Korondi, "Smartphone emotions based on human-dog interaction," in *Proceedings of International Conference on Cognitive Infocommunications*, 2011, pp. 1–6.
- [18] T. McCalley and A. Mertens, "The pet plant: Developing an inanimate emotionally interactive tool for the elderly," *Persuasive Technology, Lecture Notes in Computer Science*, vol. 4744, pp. 68–79, 2007.
- [19] F. Zhou, J. Jiao, S. Chen, and D. Zhang, "A case-driven ambient intelligence system for elderly in-home assistance applications," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 41, no. 2, pp. 179–189, 2011.
- [20] M. Mather, T. Canli, T. English, S. Whitfield, P. Wais, K. Ochsner, D. G. John, and L. L. Carstensen, "Amygdala responses to emotionally valenced stimuli in older and younger adults," *Psychological Science*, vol. 15, no. 4, pp. 259–263, 2004.
- [21] S. S. Wong, E. M. Heiby, V. A. Kameoka, and J. P. Dubanoski, "Perceived control, self-reinforcement, and depression among Asian American and Caucasian American elders," *Journal of Applied Gerontology*, vol. 18, no. 1, pp. 46–62, 1999.
- [22] J. Cohen-Mansfield, "Perceived control, reinforcement, satisfaction, and depression in community elderly," *Journal of Applied Gerontology*, vol. 9, no. 4, pp. 492–503, 1990.
- [23] P. K. Adelman, "Multiple roles and psychological well-being in a national sample of older adults," *Journal of Gerontology*, vol. 49, no. 6, pp. s277–s285, 1994.
- [24] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.
- [25] N. A. Remington, L. R. Fabrigar, and P. S. Visser, "Reexamining the circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 79, no. 2, pp. 286–300, 2000.
- [26] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 16, no. 4, pp. 261–292, 1996.
- [27] A. Mehrabian, C. Wihardja, and E. Ljuggren, "Emotional correlates of preferences for situation-activity combinations in everyday life," *Genetic, Social and General Psychology Monographs*, vol. 123, no. 4, pp. 461–477, 1997.
- [28] C. S. McCrae, J. P. H. McNamara, M. A. Rowe, J. M. Dzierzewski, J. Dirk, M. Marsiske, and J. G. Craggs, "Sleep and affect in older adults: Using multilevel modeling to examine daily associations," *Journal of Sleep Research*, vol. 17, no. 1, pp. 42–53, 2008.
- [29] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: the PANAS scales," *Journal of Personality and Social Psychology*, vol. 54, no. 6, pp. 1063–1070, 1988.
- [30] A. Mehrabian, "Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression," *Journal of Psychopathology and Behavioral Assessment*, vol. 19, no. 4, pp. 331–357, 1997.
- [31] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychological Review*, vol. 63, no. 2, pp. 81–97, 1956.
- [32] A.-H. Tan, G. A. Carpenter, and S. Grossberg, "Intelligence through interaction: Towards a unified theory for learning," in *Proceedings of International Symposium on Neural Networks, D. Liu et al. (Eds.) LNCS 4491, Part I*, 2007, pp. 1094–1103.
- [33] G. A. Carpenter and S. Grossberg, "Adaptive Resonance Theory," in *The Handbook of Brain Theory and neural Networks*. MIT Press, 2003, pp. 87–90.
- [34] D. Wang and A.-H. Tan, "Creating autonomous adaptive agent in a real-time first-person shooter computer game," *IEEE Transactions on Computational Intelligence and AI in Games*, under 2nd review.
- [35] G. A. Carpenter and S. Grossberg, "ART 2: Self-organization of stable category recognition codes for a analog input patterns," *Applied Optics*, vol. 26, no. 23, pp. 4919–4930, 1987.
- [36] G. A. Carpenter, S. Grossberg, and D. Rosen, "ART 2-A: An adaptive resonance algorithm for rapid category learning and recognition," *Neural Networks*, vol. 4, no. 4, pp. 493–504, 1991.