# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

**6,900**
Open access books available

**186,000**
International authors and editors

**200M**
Downloads

**154**
Countries delivered to

Our authors are among the

**TOP 1%**
most cited scientists

**12.2%**
Contributors from top 500 universities

CLARIVATE ANALYTICS
**BOOK CITATION INDEX**
INDEXED

**WEB OF SCIENCE**™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

## Interested in publishing with us?
## Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

# Activity Recognition Using Probabilistic Timed Automata

Lucjan Pelc[1] and Bogdan Kwolek[2]
*[1]State Higher Vocational School in Jarosław,*
*[2]Rzeszów University of Technology*
*Poland*

## 1. Introduction

Activity recognition focuses on what is happening in the scene. It endeavors to recognize the actions and goals of one or more actors from a sequence of observations both on the actor actions and the environmental conditions. Automated recognition of human activity is essential ability that may be used in the surveillance to provide security in indoor as well as outdoor environments. Understanding human activity is also important for human-computer-interaction systems including tele-conferencing and for content-based retrieval of video from digital repositories.

The main technique utilized in activity recognition is computer vision. In vision-based activity recognition, a great deal of work has already been done. This is partially due to increasing computational power that allows huge amount of video to be processed and stored, but also due to the large number of potential applications. In vision-based activity recognition, we can distinguish four steps, namely human detection, human tracking, human activity recognition and then a high-level activity evaluation.

A method of (Viola et al., 2003) detects a moving pedestrian in a temporal sequence of images. A linear combination of filters is applied to compute motion and appearance features that are then summed to determine a cumulative score, employed afterwards in a classification of the detection window as including the moving object. For vision based activity recognition, tracking is the fundamental component. The entity must be first tracked before the recognition can take place. Briefly, the goal of visual tracking is to find and describe the relative position change of the moving object from one frame to another in the whole sequence, while the task of action recognition is to classify the person's action given the person's location, recent appearance, etc. Kalman filters (Crowley & Berard, 1997; Kwolek, 2003) and particle filtering–based algorithms (Nait-Charif & McKenna, 2003) are utilized extensively for object tracking in this domain. These algorithms generally involve an object state transition model and an observation model, which reflect both motion and appearance of the object (Haykin & de Freitas, 2004). After tracking of the moving objects the action recognition stage occurs, where Dynamic Time Warping (Myers et al., 1980; Myers & Rabiner, 1981) and Hidden Markov Models (Brand & Kettnaker, 2000) are employed very often at this stage. Sophisticated stochastic models such as Dynamic Bayesian Networks (Albrecht et al., 1997; Ghahramani, 1997), Stochastic Context Free

Grammar (Pynadath et al., 1998), Probabilistic State Dependent Grammars (Pynadath et al., 2000), Abstract Hidden Markow Models (Bui et al., 2002), among others, were elaborated in order to represent high-level behaviors.

In this chapter, we focus on recognition of student activities during a computer-based examination where some knowledge about the layout of the scene is known. One characteristic of such activities is that they exhibit some specific motion patterns. The recognition is done on the basis of coordinates of the tracked heads, the activated activity areas and the probabilistic timed automata.

## 2. Relevant work

In the past decade, there has been intensive research in designing algorithms for tracking humans and recognizing their actions. An overview of work related to modeling and recognizing people's behaviors, particularly largely structured behaviors, can be found in work (Aggarwal & Cai, 1999). A more recent survey on recognizing of behaviors in surveillance images can be found in (Hu et al., 2004). There is now a rich literature on vision based action recognition. In this section, we focus on approaches and applications that are closely related to our work.

In work of (Rota & Thonnat, 2003), the video interpretation encompasses incremental recognition of scene states, scenarios and behaviors, which are described in a declarative manner. A classical constraint satisfaction algorithm called Arc Consistency-4 or AC4, is utilized to reduce the computation time for the process of recognizing such activities. The system described in work (Madabhushi & Aggarwal, 2000) is capable to recognize activities using head movement. The system is able to recognize 12 activities based on nearest neighbor classification. The activities include: standing up, sitting down, bending down, getting up, etc. A recognition rate about of 80% has been reported in this work.

The Finite State Machine (FSM) to model high-level activities has been used in work (Ayers & Shah, 2001). However, the approach presented in the mentioned work does not account for uncertainty in the model. State machine-based representations of behaviors have also been utilized in work (Bremond & Medioni, 1998), where deterministic automata in order to recognize airborne surveillance scenarios with vehicle behaviors in aerial imagery have been employed. Non-deterministic finite automaton has been employed in work (Wada & Matsuyama, 2000) as a sequence analyzer. An approach for multi-object activity recognition based on activity driven selective attention has been proposed. Bayesian networks and probabilistic finite-state automata were used to describe single-actor activities in work (Hongeng et al. 2004). The activities are recognized on the basis of the characteristics of the trajectory and the shape of the moving blob of the actor. The interaction between multiple actors was modeled by an event graph.

Recognition of mutual interactions between two pedestrians at blob level has been described in work (Sato & Aggarval, 2004). Most of the research connected with recognition of human interactions considers multiple-person interactions in remote scenes at a coarse level, where each person is represented as a single moving box. An extension of Hidden Markov Models, called Behavior Hidden Markov Models (BHMMs) has been presented in work (Han & Veloso, 1999) in order to describe behaviors and interactions in a robot system. Using such a representation an algorithm for automatically recognizing behaviors of single robots has been described too.

Hidden Markow Models (HMMs) are popular state-based models. In practice, only the observation sequence is known, while the underlying state sequence is hidden, which is why they are called Hidden Markov Models. HMMs have been widely employed to represent temporal trajectories and they are especially known for their application in temporal pattern recognition. A HMM is a kind of stochastic state machines (Brand et al., 1997), which changes its state once *every* time unit. However, unlike finite state machines, they are not deterministic. A finite state machine emits a deterministic symbol in a given state. It then deterministically transitions to another state. HMMs do neither deterministically, rather they both transition and emit under a probabilistic model. Its use consists in two stages, namely, training and recognition. HMM training stage involves maximizing the observed probabilities for examples belonging to a class. In the recognition stage, the probability with which a particular HMM emits the test symbol sequence corresponding to the observations is computed. However, the amount of data that is required to train a HMM is typically very large. In addition, the number of states and transitions can be found using a guess or trial and error and in particular, there is no general way to determine this. Furthermore, the states and transitions depend on the class being learnt. Despite such shortcomings the HMMs are one of the most popular algorithms employed in recognition of actions.

In our previous work related to action recognition, we presented a timed automata based approach for recognition of actions in meeting videos (Pelc & Kwolek, 2006). Timed automata are finite state machines extended about possibility for modelling of the behavior of real-time systems over time (Alur & Dill, 1994). A declarative knowledge provided graphically by the user together with person positions extracted by a tracking algorithm were used to generate the data for recognition of actions. The actions were formally specified as well as recognized using the timed automata.

In this chapter, we present a system for recognition of high-level behaviors of people in complex laboratory environments. The novelty of the presented approach is in the use of probabilistic timed automata (PTA). The probabilistic timed automata can model state-dependent behaviors, and with the support of time, probabilistic inference of high-level behaviors from low-level data. The PTA-based recognition module of behaviors takes sequences of coordinates of observed heads that are determined by the tracking module. Some declarative knowledge that has been specified graphically in advance by the system supervisor together with such coordinates is utilized to prepare the input data for the automata recognizing behaviors under uncertainty. The system also recognizes person-to-person interactions, which in our student examination scenario are perceived as not allowed behaviors.

## 3. Vision-based person tracking

Vision-based recognition of human activities involves extraction of relevant visual information, representation that information from the point of view of learning and recognition, and finally interpretation and evaluation of activities to be recognized. Image sequences consist of huge quantity of data in which the most relevant information for activity recognition is contained. Thus, the first step in activity recognition is to extract the relevant information in the form of movement primitives. Typically, this is achieved through vision-based object detection and tracking.

Tracking and activity recognition are closely related problems. A time series, which has been extracted by an object tracker provides a descriptor that can be used in a general

recognition framework. Robust detection and tracking of moving objects from an image sequence is a substantial key for reliable activity recognition. Much tracking methods can be applied in scenarios with simple backgrounds and constant lighting conditions. Unfortunately, in real scenarios only occasionally do such situations arise. Typically, tracking requires consideration of complicated environments with difficult visual scenarios, under varying lighting conditions.

The shape of the head is one of the most effortlessly recognizable human parts and can be sufficiently well approximated by an ellipse. Its shape undergoes relatively little changes in comparison to changes of the human silhouette. In our scenario the position of the head is very useful because on the basis of its location we can recognize the actions consisting in looking at the terminal of a neighboring student. Moreover, on the basis of the location of the head we can determine the person's movement through the scene and in consequence we can recognize several actions like entering the scene, leaving the scene, standing up, sitting down, using a computer terminal, and so on.

The participant undergoing tracking can make rotations of both his/her body and head and thus the actions should be identified in either the frontal and lateral view. This implies that the usage of only color information for person tracking in long image sequences can be infeasible. In work (Kwolek, 2004) it has been demonstrated a tracker that has proven to be very useful in long-term tracking of people attending a meeting. This particle filter based tracker is built on gradient, color and stereovision. Human face is rich both in details and texture and consequently the depth map covering a face region is usually dense. The algorithm can track a person's head with no user intervention required. More importantly, this algorithm is efficient enough to allow real-time tracking on typical 850 MHz personal computer with PIII. It can accurately track in real-time multiple subjects in most situations. The detection of person entrance has also been done on the basis of the head. The entering and leaving the scene by participants of the exam is detected in entry and exit zones on the basis of method described in (Kwolek, 2005). Assuming that the person's head is relatively flat and that the entrance should be done at some distance from the camera we can suppress pixels not belonging to the person.

## 4. Activity recognition using probabilistic timed automata

### 4.1 The problem

The aim of the system is to recognize activities as well as to detect abnormal activities (suspicious and forbidden) that can take place during examination of the students. During the exam the students solve individually some tasks using computers and the collaboration between students is not permitted. In other words, each student should solve his/her task one-self, without looking into the computer screen of the neighbor. During the unaided work the student must not change workplace and take an empty workplace, and particularly, crib another student's solution from the computer screen, if such a student temporally left his/her workplace in order to pass the oral part of the exam in another part of the laboratory or lecturer's room. Additionally, the system should recognize the start as well as the end of the activities in order to record the corresponding key-frames.

Figure 1 depicts a scene that has been shot in a typical laboratory environment. The rectangles that are overlaid on the image are employed in detection of activity areas in order to pre-segment low-level data for recognition. In work (Pelc & Kwolek, 2006) the timed automata were used in action recognition and a person was required to continuously

occupy positions within rectangular activity areas for specified in advance time intervals. It this work, the probabilistic timed automata are employed to model state-dependent activities, and with the support of time, to perform probabilistic inference of high-level behaviors from low-level data. The activity recognition under uncertainty is done on the basis of sequences of head coordinates. Some declarative knowledge in form of rectangular areas that can be activated together with such coordinates is utilized to pre-segment the input data for the automata. The system recognizes also person-to-person interactions, which in our student examination scenario are perceived as not allowed behaviors.
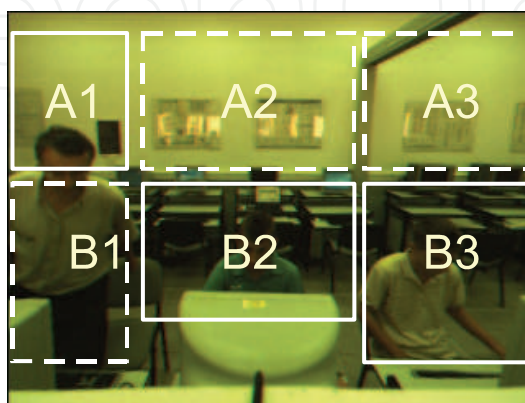


Fig. 1. A view of the scene from the camera

Using probabilistic timed automata the system recognizes the following activities:

- A_Work – work of the examined student,
- A_PartialC – partial collaboration of two students,
- A_FullC – full collaboration of two students,
- A_Suspicious – suspicious action of a student,
- A_Forbidden – forbidden action of a student.

As we already mentioned above, the system determines also both the start and the end of such activities in order to record some important key-frames. A more precise explanation of actions to be recognized will be given in the remaining subsections in coherence with description of automata.

### 4.2 Probabilistic timed automata

A description of probabilistic timed automata can be found in (Alur and Dill, 1994; Kwiatkowska et al., 2004). It is worth to note that the probabilistic timed automaton can be represented via a directed graph. In such a graph-based representation the nodes stand for the states, whereas the edges are labeled by actions. Time domain is represented by positive integer values. The variables of such a type are called clocks. They are employed to formulate state invariants and guards for the edges. After transition along an edge the clocks can be reset. In order to express probabilities the edges can be forked and they can connect more than two states. The state that will be reached by the automata after performing the action connected with the edges depends on the probability (assuming positive real values) of the given transition. Figure 2. depicts an example of elementary probabilistic timed automaton. This automaton consists of:

- two states: State0 and State1, and one clock $t$,
- two state invariants: $t \leq T$ and true, respectively, where $T$ is some integer parameter assuming positive values, whereas true means that the invariant is always satisfied,

- one action: a,
- one guard of the action: $t \geq T$,
- reset function of the clock, $t:=0$,
- probabilities of the transitions: $p$ and $1-p$.

At the beginning the automaton is in the state State0 and it can occupy this state for a period of time not longer than $T$. If in a period of time $T$ an action a occurs, then automaton may change its state. The moment of the transition is determined by the guard $t \geq T$, which together with the initial state invariant determines the moment of the transition. The automaton makes the transition that has been caused by the action a and resets the clock $t$. The automaton may transit to State1, what occurs with probability equal to $p$ or remain in the state State0, what occurs with the probability $1-p$.
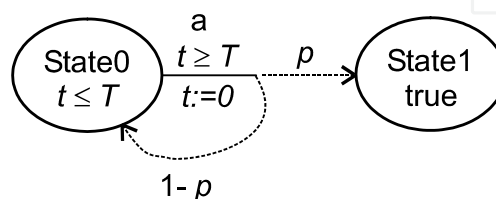


Fig. 2. Example of an elementary probabilistic timed automaton

It is worth to note that the passage of time is an action itself.

## 4.3 PTA-based action recognition

The PTA-based action recognition is done on the basis of four different concurrent automata, which synchronize themselves if needed. The aim of the automata is to recognize the following actions:

- PTA_ElementaryActionPx - recognition of an elementary action of the person Px,
- PTA_ActionWorkPx – recognition of action A_Work of person Px,
- PTA_Interact – recognition of collaboration between persons and recognition of actions not allowed: A_PartialC, A_FullC, A_Suspicious, A_Forbidden,
- PTA_StartEnd – determination of the start, the end as well as persistence of the actions undergoing recognition.

For each person the system activates separate instances of the automaton PTA__ElementaryActionPx, where x assumes values 1, 2 and 3.

In description of the PTA-based action recognition that follows in four subsequent subsections we utilize the following notation:

- names of the automata begin with PTA_, for example PTA_Interact,
- actions recognized by the system begin with A_, for example, action work is denoted as A_Work,
- actions moving an automaton from one state to another are denoted by names, which begin with a_, for example, the action connected with a step of computation is denoted by a_step,
- names of the states of the automata begin with S_, for example, in case of recognition of the action A_Work the state of the automaton PTA_ActionWorkPx is denoted as S_Work.

The system recognizes 48 actions, including actions listed in Section 4.2. Such a set of actions consists of 8 basic actions, 8 elementary actions and 4 auxiliary actions for each basic action, see Table 1.

| Type of action | Denotation | Description |
|---|---|---|
| **Basic actions** | A_NWork | no work |
| | A_PWork | probable work |
| | A_Work | work |
| | A_NoC | no collaboration |
| | A_PartialC | partial collaboration |
| | A_FullC | full collaboration |
| | A_Suspicious | suspicious action |
| | A_Forbidden | forbidden action |
| **Elementary actions** | A_Desk1, ..., A_Desk3 | Standing at workplace 1, 2, 3, respectively |
| | A_Outside | Person is outside of the camera's field of view |
| | A_Chair1, ..., A_Chair3 | Seating at workplace 1, 2, 3, respectively |
| | A_Move | Moving between workplaces |
| **Auxiliary actions** | A_Start_A_NWork, ... A_Start_A_Forbidden | Start of basic action A_NWork, ..., A_Forbidden, respectively |
| | A_During_A_NWork, ... A_During_A_Forbidden | Persistence of basic action A_NWork, ..., A_Forbidden, respectively |
| | A_End_A_NWork, ... A_End_A_Forbidden | End of basic action A_NWork, ..., A_Forbidden, respectively |
| | A_No_A_NWork, ... A_No_A_Forbidden | Lack of basic action A_NWork, ..., A_Forbidden, respectively |

Table 1. Statement of actions that are recognized by the system

### 4.3.1 Automaton PTA_ElementaryActionPx

The automaton PTA_ElementaryActionPx is designated for recogniton of elementary actions, which are listed in Table 1. It is depicted in a simplified form in Fig. 3.

The detection of an elementary action is connected with reaching by the automaton of the suitable state. The states, except the states S_Outside and S_Move, are connected with the presence of the person in the suitable rectangular area depicted in Fig. 1. Table 2 lists the connections between the active areas, the states of the automaton PTA_ElementaryActionPx, and the recognized elementary actions.

The state S_Move reflects the absence of a person in the assumed activity areas of the scene, see Fig. 1. In such a situation we assume that a person moves between the workplaces. The state S_Outside is reached when a person is out of the camera's field of view.

From the analysis of the automaton shown in Fig. 3 we can notice that the PTA-based activity recognizer detects entry and exit to/from the activity areas of the scene depicted in Fig. 1. For example, the entry event into the area A1 in the Fig. 1, and then exit from this area is connected with action (edge) a_dk1_on, which transits the automaton to state S_Desk1,

and after that the action a_dk1_off deriving the automaton from that state, see Fig. 3. Analogously, for the area A2 it would be a_dk2_on and a_dk2_off, whereas for the area B1 (S_Chair1 in automaton) a_ch1_on and a_ch1_off, and so on.



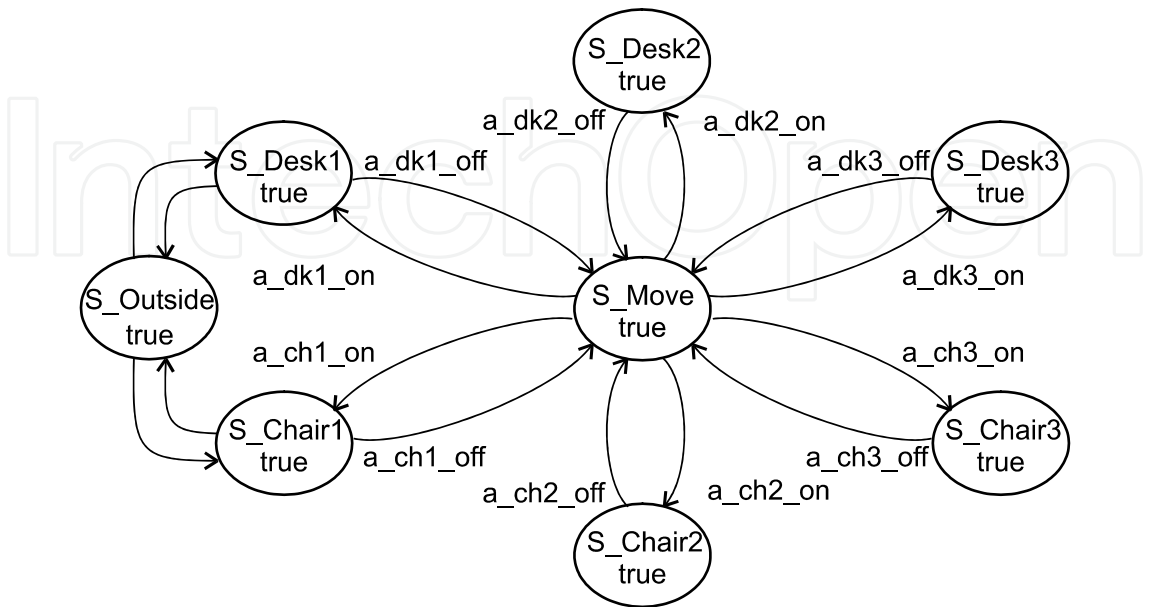Fig. 3. The automaton for recognition of elementary actions

| Area of scene | State of the automaton PTA_ElementaryActionPx | The recognized elementary actions |
|---|---|---|
| A1 | S_Desk1 | A_Desk1 |
| A2 | S_Desk2 | A_Desk2 |
| A3 | S_Desk3 | A_Desk3 |
| B1 | S_Chair1 | A_Chair1 |
| B2 | S_Chair2 | A_Chair2 |
| B3 | S_Chair3 | A_Chair3 |

Table 2. Connections between the active areas, the states of the automaton PTA_ElementaryActionPx, and the recognized elementary actions

### 4.3.2 Automaton PTA_ActionWorkPx

The action A_Work can be recognized with some likelihood. The system activates separate instances of the automaton PTA_ActionWorkPx for each person Px. Let us consider the person P1 acting in front of the area B1. The smaller number of times the area B1 has been left by the person in time *Tw*, and at the same time, the smaller was his/her total time of staying outside B1, the greater the likelihood is, that the person P1 is working. Time *Tw* stands for minimal time that should elapse between the entry and the exit of the workplace to indicate that activity relying on work has been started.

The automaton PTA_ActionWorkPx is shown at Fig. 4 (for clarity of presentation some less important details have been omitted). It consists of the following states:

S_NWork – the considered person is not working,

S_PWork – it is possible that the considered person is working,

S_Work – the considered person is working.

In case the automaton PTA_ActionWorkPx has synchronized with the suitable automata, then attaining by it the state S_NWork is equivalent with recognition by the system the action A_NWork. Analogous relationship is between S_PWork and A_PWork, S_Work and A_Work, see also Tab. 1.
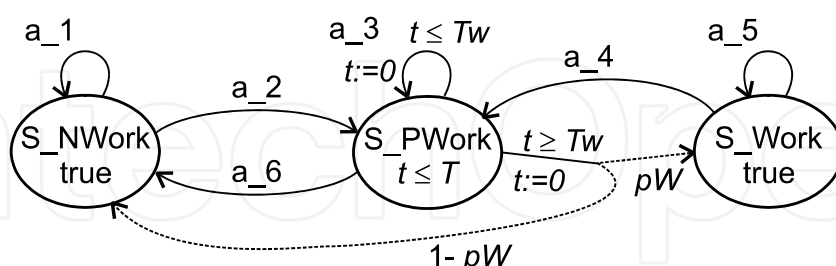


Fig. 4. Automaton PTA_ActionWorkPx

Let us assume, that the person P1 is doing alternately the actions A_Chair1 (area B1 in Fig. 1) and A_Move (outside B1). Such a probable work will be recognized as action A_PWork. In dependence of how the transitions frequent are, as well as how long they appeared in total according to time $Tw$, the person P1 may in consequence either work with the probability $pW$, what exemplifies the action A_Work, or may not work, what exemplifies the action A_NWork with probability $1\text{-}pW$.

The probability $pW$ has been determined as follows:

$$pW = w \cdot pWt + (1 - w) \cdot pWn \tag{1}$$

where $w$ denotes a weight that can assume values between 0 and 1.0, $pWt$ stands for probability resulting from the summed time of temporal absences in the area Bx, see Fig. 1, $pWn$ is a probability resulting from the number of the mentioned above absences in period of time $Tw$.

Figure 5a illustrates the method of determining the probability $pWt$. Let us denote by $tn$ the summed time of the absence in the period of time $Tw$. Then

$$pWt = \begin{cases} 1\text{-}(1/Tmax) \cdot tn & \text{for } 0 \le tn \le Tmax \\ 0 & \text{for } tn > Tmax \end{cases} \tag{2}$$

where $Tmax$ denotes maximal admissible sum of temporal absences at the workplace. The value of time $Tmax$ must be less than the value $Tw$.
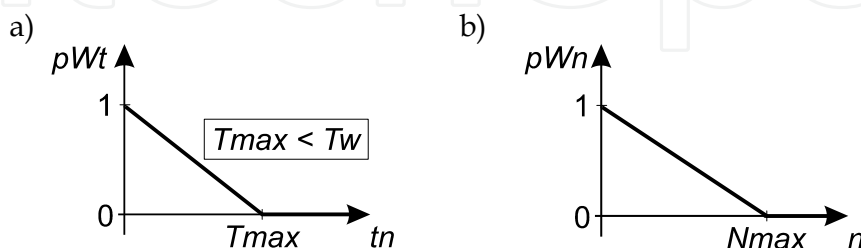


Fig. 5. The probability functions: a) $pWt$, b) $pWn$

Figure 5b illustrates the method of computation of the probability $pWn$. Let us denote by $n$ the number of absences at workplace in the period of time $Tw$. Then

$$pWn = \begin{cases} 1\text{-}(1/Nmax) \cdot n & \text{for } 0 \le n \le Nmax \\ 0 & \text{for } n > Nmax \end{cases} \tag{3}$$

where *Nmax* stands for maximal admissible number of absences in the workplace during time period equal to *Tw*.

In the sequel we explain actions of the automaton PTA_ActionWorkPx in the context of the remaining automata. The meaning of actions a_1, ..., a_6 depicted in Fig. 4, under assumption that they were done by person P1, is as follows:

- a_1 – will occur, if at the moment of occurrence of action a_step the automaton PTA_ElementaryActionPx is in other state than S_Chair1, what means that the considered person P1 is not seating at his/her workplace,
- a_2 – will occur, if at the moment of occurrence of action a_step the automaton PTA_ElementaryActionPx is in state S_Chair1, what means that the considered person is seating at his/her workplace,
- a_3 – will take place, if at the moment of occurrence of action a_step the automaton PTA_ElementaryActionPx is in the state S_Chair1 or S_Move, in other words, the considered person seats at his/her workplace, and from time to time the person lefts it temporary. Because of the guard *t<Tw*, the mentioned behavior can only take time shorter than *Tw*.
- a_4 – analogously to the action a_1, likewise, the action a_5 is analogous to action a_2,
- a_6 – will occur, if the action a_step is occurring and the automaton PTA_ElementaryActionPx is not neither in the state S_Chair1 nor S_Move,
- *t≥Tw* – action connected with passage of time.

### 4.3.3 Automaton PTA_Interact

As we already mentioned in Section 4.1 the system should recognize the activity A_Work as well as the following activities:

- Partial collaboration – A_PartialC,
- Full collaboration – A_FullC,
- Suspicious action – A_Suspicious,
- Forbidden action – A_Forbidden.

Below is a more detailed description of assumed conditions that are used in process of the recognition of the listed above activities.

1. If two persons work, then with the probability of 10% they collaborate each other and in consequence the action A_PartialC takes place.
2. If A_PartialC took place and the considered persons still work together, then the probability of persistence of the collaboration is 50%.
3. If the neighboring persons collaborate each other and one of them looks at the screen of one of the neighbors, then with the probability of 90% the action A_FullC occurs.
4. If a person is not present at his/her workplace, and another person Px is present at his/her workplace, then Px makes the suspicious action denoted as A_Suspicious.
5. If the action A_Suspicious extends in time, for time not longer than some assumed value *ta*, then the person Px makes forbidden action denoted as A_Forbidden.

Figure 6 illustrates a part of the automaton PTA_Interact for recognition of collaboration between P1 and P2. The state S_NoC stands for no collaboration.
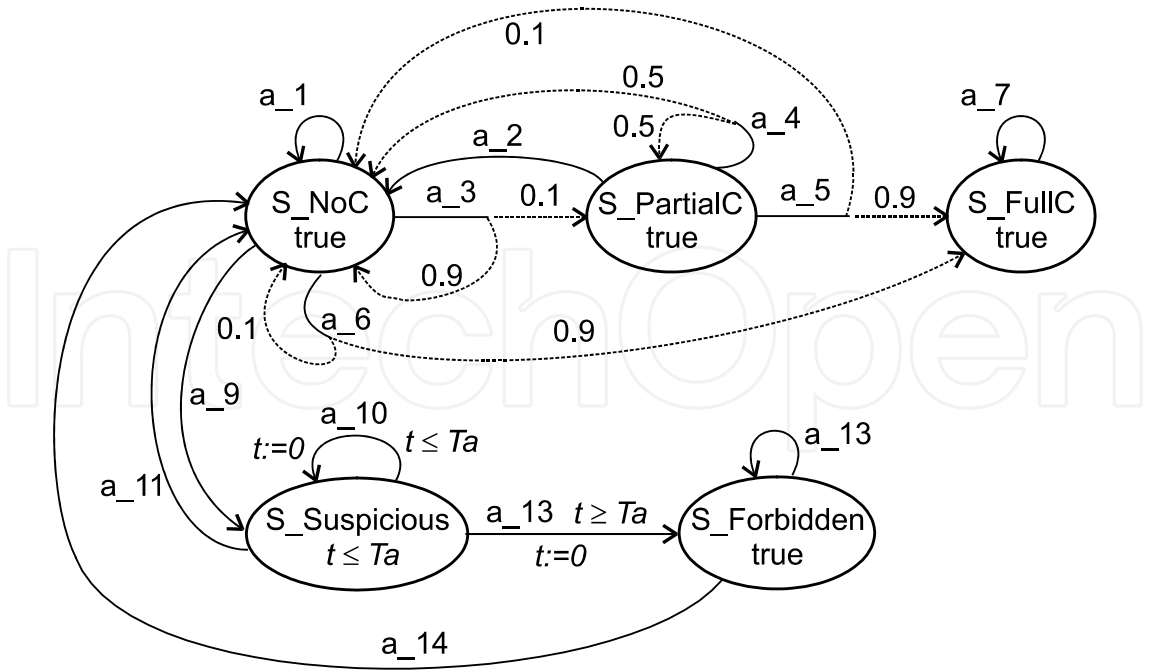
Fig. 6. Automaton PTA_Interact

The meaning of actions depicted in Fig. 6 is as follows:

- a_1 – the automata PTA_ActionWorkP1 and PTA_ActionWorkP2 are not simultaneously in the states S_Work,
- a_2, a_8 – similar meaning with a_1,
- a_3 – the automata PTA_ActionWorkP1 and PTA_ActionWorkP2 are simultaneously in the states S_Work,
- a_4, a_7 – similar meaning with a_3,
- a_5 – the automata PTA_ActionWorkP1 and PTA_ActionWorkP2 are simultaneously in the states S_Work, and additionally, the automaton PTA_ElementaryActionP1 is in the state S_Char2 or the automaton PTA_ElementaryActionP2 is in the state S_Char1,
- a_6 – similar meaning with a_5,
- a_9 – the automaton PTA_ElementaryActionP1 is in the state S_Char2, whereas the automaton PTA_ElementaryActionP2 is in the state S_Outside,
- a_10 – similarly to a_9, but the time of being in the state S_Suspicious is shorter than an assumed value *Ta*,
- a_11, a_14 – have a sense similar with the negation of the action a_9,
- a_12 – similarly to a_10, but time *Ta* elapsed,
- a_13 – similar meaning with a_9.

For the remaining persons the rest of the automaton will be analogous.

### 4.3.4 Automaton PTA_SartEnd

The automaton consists of four states for each basic action recognized by the system, see also Tab. 1. The first state is reached when the suitable action begins, the second one is when the action is in progress, the third one is reached when the action ends, whereas the fourth state exemplifies no action. The transitions between the mentioned states depend on the suitable states. The discussed automaton is employed in recording of the key-frames of the video.

### 4.4 Realization of the activity recognizer in PRISM

PRISM (Kwiatkowska et al., 2006) is a verifier of probabilistic, stochastic and non-deterministic automata. This tool has no built-in support for expressing the elapse of time in the probabilistic automata, i.e. it has no support for constructing probabilistic timed automata. In our PTA-based based system for activity recognition an event-based approach has been used to express the elapse of time. This allowed us to code the reachability graphs of our PTA-based activity recognizer for proper period of time. Such PRISM-based realization of the recognizer has been tested both on real data, i.e. from our tracker and on data obtained form a module modeling selected person activities. The module responsible for modeling the person activities is described below.

### 4.5 Simulations

For practical test of the formal model as well as simulation purposes we prepared models of the activities to be recognized. Such models of the person activities have a form of deterministic timed automata. The aim of the automata is to express real activities that can happen in our scenario. They are employed to generate the input data for the probabilistic timed automata responsible for activity recognition. Such an approach permits the analysis of the correctness of the recognition for typical activities. Furthermore, the simulation was helpful in tuning of the system through setting the values of parameters. In particular, owing to simulation experiments we chosen the following values of the parameters:

- the number and location of the activity areas,
- $Tmax$ - maximal admissible sum of temporal absences at the workplace,
- $Tw$ - minimal time of being in workplace, that allows us to draw a conclusion that the person might begin his/her work,
- $Nmax$ - maximal admissible number of absences in the workplace during time period equal to $Tw$,
- $w$ – weight in (1),
- $Ta$ – minimal admissible time of being in another than the assigned workplace.

Our simulation experiments demonstrated that practical difficulties related to too small activity areas can be compensated through larger values of $Nmax$ and smaller values of the parameter $w$. Good choice of the mentioned parameters guarantees high recognition rate despite smaller activity areas. In case of not suspicious actions, a decrease of the size of activity areas leads to smaller efficiency of the recognition.

## 5. Experiments

The system for activity recognition has been tested both on simulated as well as real data that were provided by the person tracking module. The experiments demonstrated that the system correctly recognizes the activities described previously. Figure 7 shows example images with visualization of the recognition process. The dotted line illustrates the activity zones in which no person attendance (no head) has been recognized, solid line marks the activity zones in which presence of a person has been recognized, double line informs about attendance of two persons in a single activity zone, whereas the crossed zone notifies about a forbidden action.

At Fig. 7a we can observe person P1 inclining to person's P2 side. Because this activity persisted over sufficiently long time, the system recognized the full collaboration A_FullC between the person P1 and the person P2, what has been depicted via double line overlaid on the zone B2. At Fig. 7b a full collaboration between P2 and P3 has been depicted in a similar manner. Figure 7c illustrates a situation where the person P3, who should occupy only the workplace number 3 in the zone B3, changed workplace during absence of the person P2 and now occupies his/her workplace in zone B2. Because this action was sufficiently long the system recognized the action A_Forbidden, what has been visualized via crossing the zone B2.

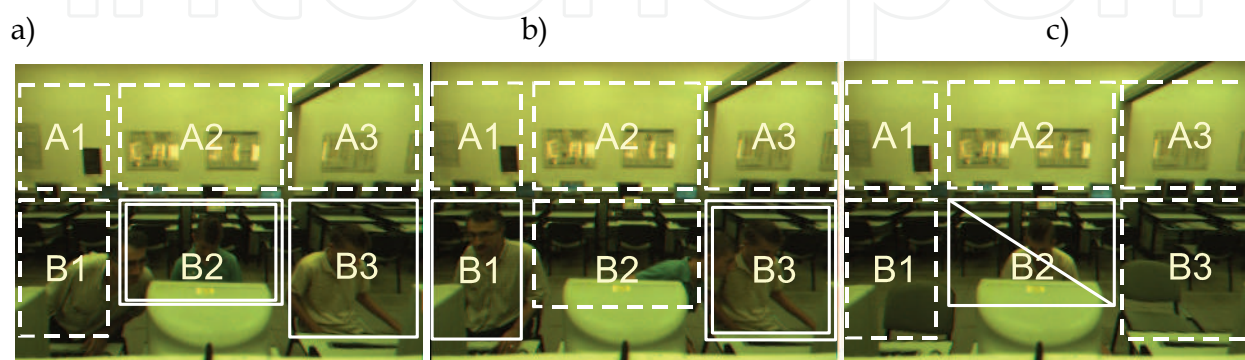a)                                        b)                                        c)



Fig. 7. Recognition of student activities during an exam, action A_FullC between P1 and P2 (a), and P2 and P3 (b), action A_Forbidden done by person P3 at workplace number 2 (c)

The images used in this work are 320 x 240 in size. They were acquired by commercial binocular Megapixel Stereo Head (Konolige, 1997). The stereo head is designed to operate with SRI's stereo engine for fast range determination on standard PC hardware. It delivers range and color images at frame rates 30 Hz with a PIII 750 MHz. Range is interpolated to 1/16 pixel, texture and consistency filters eliminate ambiguous areas in the range images. We decided to utilize stereo information in our head detection and tracking algorithm because of its usefulness in varying illumination conditions. Using it in the tracking algorithm we determine the length of the ellipse's minor axis approximating the oval shape of the head. The stereo information serves also as constraint region both in head detection and tracking.

By employing shape, color, stereovision as well as elliptical shape features our particle filter based tracker (Kwolek, 2004) can estimate the location of the person's head reliably. It is able to track the heads of persons over the whole image sequences of typical exam. Its robustness with respect to full 360-degree out-of-plane rotation, considerable head tilting, substantial but short in time head occlusions, moving people in background, as well as varying illumination conditions play very important role in the process of activity recognition without user intervention.

To test the activity recognition system, we have conducted several experiments in circumstances occurring naturally in laboratory. The aim of the experiments was to automatically detect the entrance of students into the exam, and then to track the heads. The videos as well as the 3D locations of the heads with the corresponding ellipse sizes were stored in files for further analysis. The activity recognition as well as visualization are done on the basis of such data.

## 6. Conclusions

We have developed a system for recognizing and monitoring human behaviors. It recognizes both single- and multiple actor activities. The approach is based upon probabilistic timed automata. The input data for the probabilistic timed automata are extracted via person's head tracker. By employing depth, color, as well as elliptical shape features the utilized particle filter tracks a head over a sequence of images and generates the trajectories of the head. Experimental results demonstrate the ability of the system to provide monitoring of high level behaviors in the student exam scenario. Given workplace - specific constraints, the system enabled the actions of working, collaborating, doing forbidden actions to be recognized. Although we only showed example in the context of the student exam, our system is capable to monitor a wide variety of events. One of the advantages of the presented approach is that it does not require a large amount of training data for recognition of activities at acceptable level.

## 7. References

Aggarwal, J. K., Cai, Q. (1999). Human motion analysis: A review, *Computer Vision and Image Understanding*, vol. 73, no. 3., pp. 428-440.

Albrecht, D., Zukerman, I., Nicholson, A. (1997). Bayesian models for keyhole plan recognition in an adventure game. *User Modelling and User-Adapted Interaction*, vol. 8, no. 1–2, pp. 5–47.

Alur, R., Dill, D. (1994). A theory of timed automata, *Theoretical Computer Science*, vol. 126, no. 2, pp. 183-235.

Ayers, D., Shah, M. (2001). Monitoring human behavior from video taken in an office environment, *Image and Vision Computing*, vol. 19, no. 12, pp. 833-846.

Brand, M., Oliver, N., Pentland, A. (1997). Coupled hidden Markov models for complex action recognition, In Proc. IEEE Conf. on Computer Vision and Pattern Rec., pp. 994-999.

Brand, M., Kettnaker, V. (2000). Discovery and segmentation of activities in video, *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 22, pp. 844-851.

Bremond, F., Medioni, G. (1998). Scenario recognition in airborne video imagery, In Proc. Int. Workshop on Interpretation of Visual Motion, pp. 57-64.

Bui, H. H., Venkatesh, S., West, G. (2002). Policy recognition in the Abstract Hidden Markov Model, *Journal of Artificial Intelligence Research*, vol. 17, pp. 451–499.

Crowley, J. L., Berard, F. (1997). Multi-modal tracking of faces for video communications, IEEE Conf. on Comp. Vision and Pattern Recognition, Puerto Rico, pp. 640-645.

Ghahramani, Z. (1997). Learning Dynamic Bayesian Networks, Lecture Notes in Computer Science, vol. 1387, pp. 168-197.

Han, K., Veloso, M. (2000). Automated robot behavior recognition applied to robotic soccer. In: J. Hollerbach and D. Koditschek (eds.), Robotics Research: the Ninth Int. Symposium, Springer-Verlag, pp. 199-204.

Hongeng, S., Nevatia, R., Bremond, F. (2004). Video-based event recognition: activity representation and probabilistic recognition methods, *Computer Vision and Image Understanding*, vol. 96, no. 2, pp. 129-162.

Hu, W., Tan, T., Wang, L., Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors, *IEEE Trans. on Systems, Man, and Cybernetics*, Part C, vol. 34, no. 3, pp. 334-352.

Haykin, S., de Freitas, N. (eds.) (2004). Special Issue on Sequential State Estimation. *Proceedings of the IEEE*, vol. 92, no. 3, pp. 399-574.

Konolige, K. (1997). Small Vision System: Hardware and implementation, Proc. of Int. Symp. on Robotics Research, pp. 111-116.

Kwiatkowska, M., Norman, G., Sproston, J., Wang, F. (2004). Symbolic model checking for probabilistic timed automata, Joint Conf. on Formal Modelling and Analysis of Timed Systems (FORMATS) and Formal Techniques in Real-Time and Fault Tolerant Systems (FTRTFT), vol. 3253, Lectures Notes in Computer Science, Springer-Verlag, pp. 293-308.

Kwiatkowska, M., Norman, G., Parker, D. (2006). Performance analysis of probabilistic timed automata using digital clocks, *Formal Methods in System Design*, Springer, vol. 29, pp. 33-78.

Kwolek, B. (2003). Visual system for tracking and interpreting selected human actions, *Journal of WSCG*, vol. 11, no. 2, pp. 274-281

Kwolek, B. (2004). Stereovision-based head tracking using color and ellipse fitting in a particle filter, European Conf. on Comp. Vision, LNCS, vol. 3691, Springer, 2004, pp. 192–204.

Kwolek, B. (2005). Action recognition in meeting videos using head trajectories and fuzzy histogram, *Informatica*, Special Issue on Soft Computing in Multimedia Processing, vol. 29, pp. 281-289.

Madabhushi, A., Aggarwal, J. K. (2000). Using head movement to recognize activity, In Proc. of 15th Int. Conf. on Pattern Recognition, pp. 698-701.

Myers, C. S., Rabinier, L. R., Rosenberg, A. (1980). Performance tradeoffs in dynamic time warping algorithms for isolated word recognition, *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 28, no. 6, pp. 623-635.

Myers, C. S., Rabiner, L. R. (1981). A comparative study of several dynamic time-warping algorithms for connected word recognition, *The Bell System Technical Journal*, vol. 60, no. 7, pp. 1389-1409.

Nait-Charif, H., McKenna, S. J. (2003). Head tracking and action recognition in a smart meeting room, the IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance, Graz, Austria, pp. 24-31.

Pelc, L., Kwolek, B. (2006). Recognition of actions in meeting videos using timed-automata, *Machine Graphics and Vision*, vol. 15, no. 3-4, pp. 577-584

Pynadath, D. V., Wellman, M. P. (1998). Generalized queries on probabilistic context-free grammars. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 65–77.

Pynadath, D. V., Wellman, M. P. (2000). Probabilistic statedependent grammars for plan recognition. In Proc. of the 16th Annual Conf. on Uncertainty in Artificial Intelligence, San Francisco, CA, pp. 507–514.

Rota, N. A., Thonnat, M. (2002). Activity recognition from video sequences using declarative models, Proc. European Conf. on Artificial Intelligence, pp. 673-680.

Sato, K., Aggarval, J. K. (2004). Temporal spatio-velocity transform and its application to tracking and interaction, *Computer Vision and Image Understanding*, vol. 96, no. 2, pp. 100-128.

Viola, P., Jones, M., Snow, D. (2003). Detecting pedestrians using patterns of motion and appearance, *Int. J. of Computer Vision*, vol. 63, no. 2, pp. 153-161.

Wada, T., Matsuyama, T. (2000). Multi-object behavior recognition by event driven selective attention method, *IEEE Trans. on Pattern Anal. Machine Intelligence*, vol. 22, pp. 873-887.

**Pattern Recognition Techniques, Technology and Applications**

Edited by Peng-Yeng Yin

A wealth of advanced pattern recognition algorithms are emerging from the interdiscipline between technologies of effective visual features and the human-brain cognition process. Effective visual features are made possible through the rapid developments in appropriate sensor equipments, novel filter designs, and viable information processing architectures. While the understanding of human-brain cognition process broadens the way in which the computer can perform pattern recognition tasks. The present book is intended to collect representative researches around the globe focusing on low-level vision, filter design, features and image descriptors, data mining and analysis, and biologically inspired algorithms. The 27 chapters coved in this book disclose recent advances and new ideas in promoting the techniques, technology and applications of pattern recognition.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Lucjan Pelc and Bogdan Kwolek (2008). Activity Recognition Using Probabilistic Timed Automata, Pattern Recognition Techniques, Technology and Applications, Peng-Yeng Yin (Ed.), ISBN: 978-953-7619-24-4, InTech, Available from:

http://www.intechopen.com/books/pattern_recognition_techniques_technology_and_applications/activity_recognition_using_probabilistic_timed_automata

# INTECH
open science | open minds