LMA-based Human Behaviour Analysis Using HMM

Kamrad Khoshhal[1], Hadi Aliakbarpour[1], Kamel Mekhnacha[2], Julien Ros[2],
Joao Quintas[1] and Jorge Dias[1]

[1]Institute of Systems and Robotics – University of Coimbra – Portugal
[1]{kamrad, hadi, jquintas, jorge}@isr.uc.pt
[2]Probayes SAS - Montbonnot - France
[2]{kamel.mekhnacha,julien.ros}@probayes.com

Abstract. In this paper a new body motion-based Human Behaviour Analysing (HBA) approach is proposed for the sake of events classification. Here, the interesting events are as normal and abnormal behaviours in a Automated Teller Machine (ATM) scenario. The concept of Laban Movement Analysis (LMA), which is a known human movement analysing system, is used in order to define and extract sufficient features. A two-phase probabilistic approach have been applied to model the system's state. Firstly, a Bayesian network is used to estimate LMA-based human movement parameters. Then the sequence of the obtained LMA parameters are used as the inputs of the second phase. As the second phase, the Hidden Markov Model (HMM), which is a well-known approach to deal with the time-sequential data, is used regarding the context of the ATM scenario. The achieved results prove the eligibility and efficiency of the proposed method for the surveillance applications.

Keywords: Human Behaviour Analysing, Laban Movement Analysis, HMM and Bayesian Network.

1 Introduction

HBA is demanding for many applications such surveillance systems and home-cares. Human Movement Analysis (HMA) consider as a prerequisite for the body motion-based HBA. Having human movement's properties makes it possible to interpret human behaviours, which is a more complex task. As Bobick [4] believes, human behaviour comes from a sequence of performed human motions inside a scene. It means that previous knowledge of human motion is needed, to be able to understand human behaviours in the contex of a particular scenario. Thus it seems that there is a couple of issues, namely a sequence of movement and environment parameters that need to be investigated in order to estimate the human behaviour.

Hidden Markov Model (HMM), which is kind of Dynamic Bayesian Network (DBN) methods, is a well-known method for the purpose of analysing in such a sequential movement data and it can deal with previous knowledge dependencies. Remagnino and Jones [14] used a HMM approach to model parking lot environment behaviours. Oliver et al. in [9] defined some sequences of human motions to estimate the people behaviours.

There are many elements or parameters, which can affect human behaviour in different situations. Pentland and Liu in [11] discussed how to model the human behaviour in a driving situation. They believed that it is useful to have some dynamic models for each kind of driving such as relaxed driving, tight driving, etc. then classify the driver's behaviour by comparing it with the models. Nascimento et al. in [8] described a method for recognizing some human activities in a shopping space (e.g. entering, exiting, passing and browsing). They used human motion patterns, which were achieved from a sequence of displacements of each human's blob center.

Ryoo and Aggarwal have used the HMM for different level of human behaviour understanding: primitive and complex in 2D-base space [16], [17]. A deep contribution in the field of human-machine interaction (HMI), based on the concept of LMA, is performed by Rett & Dias in [15]. In their work a Bayesian model is defined for learning and classification. The LMA is presented as a concept to identify useful features of human movements to classify human gestures.

In this paper, by getting inspiration from the previous work of Rett and Dias in [15], the concept of LMA is used in order to define and extract sufficient features. A two-phase probabilistic approach has been applied to model the system's state (see Fig 1). Firstly, a Bayesian network is used to estimate LMA-based human movement parameters. Then a sequence of the obtained LMA parameters is used as the inputs of the second phase. As the second phase, the HMM, which is a well-known approach to deal with the time-sequential data, is used regarding the context of the ATM scenario. The achieved results prove the eligibility and efficiency of the proposed method for the surveillance applications.
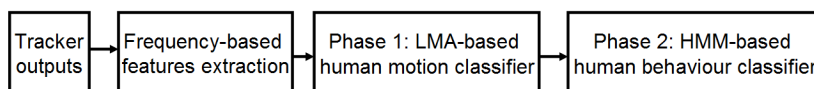
| Tracker outputs | → | Frequency-based features extraction | → | Phase 1: LMA-based human motion classifier | → | Phase 2: HMM-based human behaviour classifier |

Fig 1. The methodology diagram

This paper is arranged as following. Section 2 describes the contribution to sustainability of the paper. Section 3 presents a frequency-based feature extraction method. Then LMA concept and the human motion understanding approach are described in section 4. Section 5 describes the second phase of our HMM-based classification part, which is defined for human behavior understanding. Section 6 presents experimental part, and Section 7 closes with a conclusion and an outlook for future works.

2   Contribution to sustainability

In this paper, a new body motion-based HBA approach is proposed for the sake of events classification. The key is that a couple of classifier was used based on frequency-domain features and LMA concept. The impact of this

Fig. 2. Robberies state in an ATM scenario - PROMETHEUS dataset

paper will be a reliable behaviour Analysing system using human body motion features to analyse different human-being events. The system will be useful in many applications especially in surveillance systems and smart-homes, which are growing very fast in the world. In this paper, the interesting events are categorized as normal and abnormal behaviours in ATM scenarios (see Fig. 2).

3 Frequency-based feature extraction

We believe that the frequency-based features are suitable features to achieve the *LMA.Effort* parameters and recognize human motion, as can be seen in our previous
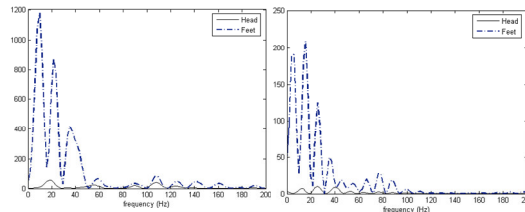

Fig. 3. PS of acceleration signals of two body parts (head and feet) for (left) running and (right) walking movement.

work [6] in detail (*LMA.Effort* parameters shall be introduced in the next section). Frequency-based features can be obtained by using Fast Fourier Transform (FFT) and Power Spectrum (PS) techniques on the input signals ([13] and [5]). The acceleration signals of the body parts are supposed to be available as the input data, and then FFT and PS of those signals are extracted as can be seen in Fig. 3.

By having all these PS signals for each selected body part acceleration signal in different actions, and collecting some coefficients (peak) of the extracted signals, like [6] which collected first four coefficients of each subdomain of PS signals, we have sufficient features in our application. Thus, four features for two parts of body are defined to be use for classification of various actions. $Max\{^{Acc}f_i^{pb}\}$ denotes the maximum content of each i subdomain-frequency of acceleration signal for each parts of body (pb). The set of pb and subdomain frequency are defined as {Head feet} and {(0-10), (11-20), (21-30), (31-40)} in Hz unit.

4 LMA-based human motion modelling

Laban Movement Analysis (LMA) is a known method for observing, describing, notating, and interpreting human movement, that was developed by Rudolf Laban, who is widely regarded as a pioneer of European modern dance and theorist of movement education [19] about 60 years ago. Norman Badler's group was the first group who attempted to re-formulate Laban framework in computational models since 1993 [1],

[19]. Recently Dias's group also had several interesting works around LMA since 2007 [15].

The theory of LMA consists of several major components, though the available literature is not in unison about their total number. The works of Norman Badlers group [19] mentioned five major components; *Body*, *Effort*, *Space*, *Shape* and *Relationship*.

One of the most important components of LMA is *Effort* that we tried to obtain it and then based on that, reach to human behaviour. *Effort* or dynamics is a system for understanding the more subtle characteristics about the way a movement is done with respect to inner intention. The difference between punching someone in anger and reaching for a glass is slight in terms of body organization - both rely on extension of the arm. The attention to the strength of the movement, the control of the movement and the timing of the movement are very different. *Effort* has four subcategories; Time, Space, Weight and Flow, and each of them has two opposite polarities which are sudden/sustained, direct/indirect, strong/light and bounded/free, respectively (for more information: [6]).

In the first phase, LMA parameters were obtained to analyse some interesing human movements depend on the scenario and tracker outputs. Based on the collected data and our tracking outputs, we could just rely on position of feet and head. Thus, we could have just acceleration signal of these body parts.

The interesting activities in ATM scenario are standing, walking, running and falling down. In our previous work [6] based on the results, we realized that by having just a couple of body parts positions insead of six parts of body, and having just a couple of states for *Effort.time* (sudden and sustained) are not enough to distinguish the interesting activities. Thus we discretized *Effort.time* to four states, as can be seen in Fig. 4. Then the outputs of this Bayesian net are all the probabilities of the interesting activities.
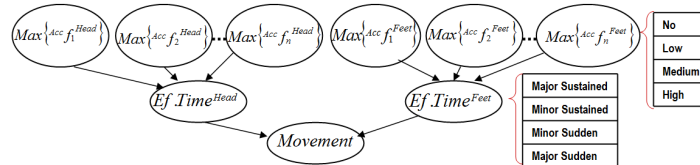

Fig. 4. LMA-based Bayesian Net

5   Concurrent HMM-based human behaviour modeling

For behaviour recognition, we are interested in detecting the current behaviour amongst *N* known behaviours (i.e. the behaviour library). For this purpose, using a concurrent HMM architecture is proposed.

5.1   Principle

A concurrent HMM is composed of several HMMs, each one describing one class (see Fig. 5 left). To summarize, the concurrent HMM centralizes the on-line update of the behaviour belief and contains:

1.   The set of HMMs representing basic behaviour library (one HMM per behaviour);

2.   The transition between behaviours model that could be either defined by hand (by an expert), or learnt from annotated data.
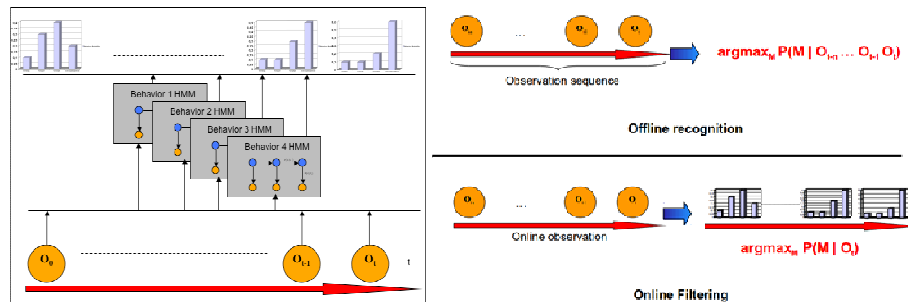

Fig. 5. Concurrent HMMs (left) Recognition Modes (right).

HMMs are used to characterize an underlying Markov chain which generates a sequence of states. The term Hidden in the HMM name comes from the fact that the sequence of states is not directly observable. Instead the states generate an observable sequence. Thus, the output depends on the current state and on previous outputs. These tools are widely used in the field of sound processing [12], gene finding and alignment in DNA sequences [10]. They were introduced by Andre Markov in [7] and developed in [2].

HMM are widely employed in the field of computer vision to recognize gesture or human behaviour [3] in these applications, the observation variables are features extracted for video data. The principle of an HMM is presented on inside of the Fig. 5 (left) (the four HMMs) in which the top (blue) circles (*S*) represent the state variables and the bottom circles (*O)* represent the observation ones in a sequence of times. In our application, each HMM describes a human behaviour and is learned using a training dataset composed of labeled observation sequences that are low-feature extracted from the LMA.

5.2   Construction/Learning

Constructing a concurrent HMM consists in:
- Learning the set of HMM models representing the behaviour library (one HMM per behaviour) using an annotated data set.
- Defining the transition matrix between the behaviours. This transition model could be either defined by hand (by an expert), or learnt from an annotated data set.

Learning the behaviour transition model is straightforward and consists in computing simple statistics (histograms) of transitions using the annotated data set. Learning the underlying HMM models (a HMM per behaviour) is more complex. It can be divided into two sub-problems:

1. Finding the optimal number of states N. The optimal number of internal states within the HMMs could be chosen by hand thanks to an expert. In this case no algorithm is needed and the learning of the HMM is reduced to the learning of its parameters. However, since an HMM is a Bayesian Network, a score that allows a compromise between fitting learning examples (D) and the ability of generalization (see the Occam Razor Principle) can be employed to find it automatically [3]. For example, the classical Bayesian Information Criterion [18] that maximizes the likelihood of the data while penalizing large size model can be used:

$$BIC(n,D) = \log(likelihood(D,n)) - \frac{1}{2} \times nparams(n) \times \log(|D|)$$

In this case, the optimal number of states is given by: $n* = \arg\max_n BIC(n, D)$

2. Learning the parameters of the HMM given N (i.e., the transition matrix $P(S_t | S_{t-1})$, the observation distribution $P(O_t | S_t)$, and the initial state distribution $P(S_0)$). The idea is find the parameters that maximize the data likelihood. For this purpose the methods generally employed are the classical EM algorithm (aka Baum-Welch algorithm in the HMM context), or the Iterative Viterbi algorithm.

5.3   Recognition

As previously emphasized, the concurrent HMM is used to recognize on-line or off-line the current behaviours amongst N known behaviours (see Fig. 5 (right)). This is easily performed by finding the HMM M that maximizes $P(M | O_{t-n},...,O_t)$ for the off-line case (or $P(M | O_t)$ for the on-line case).

6   Experimental results

In our experiments, the dataset of PROMETHEUS project has been used. Among the different various surveillance-related scenarios which exist in the database, some ATM scenarios have been selected for our intention. There is a network of cameras which observe the scene. In the ATM scenario, there are several behaviours or states which come from the involved people in the scene, such as waiting, taking money, exiting, entering and robbery at ATM area. The most interesting state in this kind of scenario usually is robbery which is an abnormal situation that can happen easily, because usually there is no any support for its security around most of the ATMs. The robbery event consists some human activities which can be defined as "when person A walks toward a person B standing close to the machine, stand by a very short time and then escape".

Using the available tracking outputs in the database, the frequency-based features are extracted by applying the proposed method in Section 3. Then the LMA-based human motion classifier, (which is a Bayesian Network [6]) is applied to the achieved frequency-based features. Then the probabilistic outputs of the LMA-based classifier are fed to the HMM as the observation data. A couple of states, namely normal and abnormal, are defined as the outputs of the HMM classifier. The abnormal state

corresponds to a situation in which a robbery is happening near the ATM and the normal state corresponds to the other activities.

In this case, the main environment parameter is the position of people related to the ATM. Thus a threshold distance for selecting the persons who are around the ATM and can involve in the scenario, is defined. By having the ATM scenarios data in PROMETHEUS dataset, we defined the robbery state as usually the robber waits in ATM's area and then goes to near of a person how is taking money from the ATM and then escapes. As can be seen in the definition, this state happens by a sequence of sub-states. Thus HMM's approach is used to model this state and normal state also. The LMA output which has four probabilities for standing, walking, running and falling down action for each person is used and collected some of those data for learning and others for classification of the HMM model.

Four scenes with different durations were collected. As we mentioned before, the interesting event which is abnormal state in this kind of scenario, is robbery. In this level, a 10 second's window on the data which will be shifted 1 second along the time is defined. By 1-second shifting the defined window along the time, 148 (windows) samples will be obtained. Between these 148 samples, there are 139 normal and 8 abnormal samples which correspond to the normal and abnormal (robbery) events, respectively. It should be mentioned that, 61 samples of normal data and 4 samples of abnormal ones have been randomly selected for learning process and the others (78 samples of normal and 4 samples of abnormal) for classification process.

Table 1. Classification results

	Normal	Robbery	%
Normal	72	6	92
Robbery	0	4	100

Table 1 presents the obtained result. It shows that to detect abnormal behaviour the collected features are appropriate and the method is very reliable, however there are some false alarm which can be reduced by using more data to learn the HMM.

7 Conclusion and future work

In this paper a novel body motion-based HBAapproach is proposed for the sake of events classification in a ATM security scenario. The concept of LMA, which is a known human movement analysing system, is used in order to define and extract sufficient features. A two-phase probabilistic approach has been applied to model the system's state. Firstly, a Bayesian network is used to estimate LMA-based human movement parameters. For obtaining the dependancies between a sequence of human motion, HMM approach was selected for learning and classification of human behaviours. The presented results are considerable in terms of detecting all abnormal behaviour, which is very important in the security scenarios. As the future work, we intend to apply this approach to other interesting scenarios such security and smart-

home scenarios. Moreover we intend to explore human-human behaviour analysis techniques based on the LMA paramateres.

References

1. N.I. Badler, C.B. Phillips, and B.L. Webber. Simulating Humans: Computer Graphics,Animation, and Control. Oxford Univ. Press, 1993.
2. L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *The Annals of Mathematical Statistics*, 1970.
3. A. Biem. A model selection criterion for classification: Application to hmm topology optimization. In *ICDAR '03*, page 104, Washington, DC, USA, 2003. IEEE Computer Society.
4. A.F. Bobick. Movement, activity and action: the role of knowledge in the perception of motion. Philosophical Trans. of the Royal Society B:Biological Sciences,352(1358):1257-1265, 1997.
5. Fangxiang Cheng, William Christmas, and Josef Kittler. Periodic human motion description for sports video databases. In Proceedings of the 17th ICPR, 2004.
6. K. Khoshhal, H. Aliakbarpour, J. Quintas, P. Drews, and J. Dias. Probabilistic LMA-based classification of human behaviour understanding using power spectrum technique. In 13th International Conference on Information Fusion2010, EICC Edinburgh, UK, July 2010.
7. A. Markov. An example of statistical investigation of the text 'eugene onegin' concerning the connection of samples in chains. In *Lecture at the, Royal Academy of Sciences, St. Petersburg.*
8. J.C. Nascimento, M. A. T. Figueiredo, and J. S. Marques. Segmentation and classification of human activities. In Int. Workshop on Human Activity Recognition and Modelling, UK, 2005.
9. Nuria M. Oliver, Barbara Rosario, and Alex Pentland. A Bayesian computer vision system for modeling human interactions.IEEE Trans. on pattern analysis and machine intelligence, 2000.
10. L. Pachter, M. Alexandersson, and S. Cawley. Applications of generalized pair hidden markov models to alignment and gene finding problems. RECOMB '05, P. 241-248, NY, USA, 2001.
11. Alex Pentland and Andrew Liu. Modeling and prediction of human behavior. IEEE Intelligent vehicles 95, pages 350_355, 1995.
12. Lawrence R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. pages 267–296, 1990.
13. H. Ragheb, S. Velastin, P. Remagnino, and T. Ellis. Human action recognition using robust power spectrum features. In 15th IEEE Int. Conf. on Image Processing, p. 753-756,2008.
14. P. Remagnino and G.A. Jones. Classifying surveillance events from attributes and behaviour. In Proceedings of the biritish machine vision conference, pages 685_694, Manchester, 2001.
15. Joerg Rett, Jorge Dias, and Juan-Manuel Ahuactzin. Laban Movement Analysis using a Bayesian model and perspective projections. Brain, Vision and AI, 2008.
16. M. S. Ryoo and J. K. Aggarwal. Recognition of composite human activities through context-free grammar based representation. In CVPR '06, p 1709-1718, 2006.
17. M. S. Ryoo and J. K. Aggarwal. Recognition of high-level group activities based on activities of individual members. In WMVC '08, p. 1-8, 2008. IEEE Computer Society.
18. G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.
19. L. Zhao and N.I. Badler. Acquiring and validating motion qualities from live limb gestures. Graphical Models, pages 1-16, 2005.