Hidden Markov Model and Its Application in Human Activity Recognition and Fall Detection: A Review

Tingting Xue$^{(\boxtimes)}$ and Hui Liu

University of Bremen, Bibliothekstraße 1, 28359 Bremen, Germany
{xue,hui.liu}@uni-bremen.de

Abstract. This paper firstly describes the research framework of Human Activity Recognition and Fall Detection, as well as Hidden Markov Model and its extension with continuous observations and hierarchical topology, namely the Continuous Density Hidden Markov Model and the Hierarchical Hidden Markov Model. Subsequently, we introduce how to apply Hidden Markov Models to the human activity modeling in Human Activity Recognition and Fall Detection based on previous literature. Finally, famous research work for smart home technology and elderly care based on Hidden Markov Models is reviewed.

Keywords: Hidden Markov model · Continuous density hidden Markov model · Hierarchical hidden Markov model · Human activity recognition · Fall detection · Activity modeling · Smart home

1 Background: Human Activity Recognition and Fall Detection

1.1 Research Framework of Human Activity Recognition (HAR) and Fall Detection (FD)

The essential research aspects in the HAR and FD study, which compose a pipeline of the research framework, are introduced as follows:

Recording Equipment and Data Acquisition. All recognition task is based on data and usually big data. So, we should first choose suitable devices and equipment for sensing the signals emitted from the human body and then use these devices to acquire enough data for the following research.

Data Segmentation and Annotation. After we have the acquired data, we need to segment them into pieces, such as single motions, activities, or activity sequences. The annotation or labeling on the segments is also required since we want to train our machine learning model with the data segments.

Digital Signal Processing (DSP) and Feature Extraction. The recorded data, or the segmented data pieces, are so-called raw data, which contain sequences of original values/vectors within a duration. In order to fit the feature extraction, DSP is a necessary step to polish these values. For example, we could apply signal normalization, amplification, or filtering to the data to make them more suitable for training and recognition. Subsequently, we can extract several features from the processed data. Simple example features of the human activity signals can be average values, variance, among others. More complex feature calculations can also be obtained directly from public feature libraries, such as [1]. So as to optimize the feature extraction and model training, sometimes feature space reduction and feature vector stacking are researched [2,3].

Human Activity Modeling of Machine Learning (ML). After preparing the data for the recognition system, we need to have a machine learning model for training and decoding. Here we can apply, for example, artificial neural networks (ANNs) [4]. As a further deepening of ANNs, Deep Neural Networks (DNNs) [5] play a cornerstone role in image recognition. On this basis, many research works have shown the ability of Convolutional Neural Networks (CNN) for image-based HAR [6]. Recently, Residual Network (ResNet) [7] shows its ability of HAR [8]. Moreover, Hidden Markov Model [9,10] is another useful human activity modeling tool that may better explain the internal structure of human activities than neural networks [11], which is the focus of this paper.

Model Training, Evaluation, Tuning and Model Application. We apply the ML model and the labeled data to train the created models, recognize new input data, evaluate the recognition system, and tune parameters and models based on the experimental analysis. If the recognition system works well for our purpose, it could be put into use for real-world application.

1.2   Sensing Technology for HAR and FD

There are many types of sensing technologies for HAR and FD in a home environment, as categorized in Fig. 1.

Basically, sensing technologies can be divided into two approaches, namely external and internal sensing methods.

In external sensing, the devices are fixed in predetermined points of interest, so the inference of activities entirely depends on the voluntary interaction of the users with the sensors. For example, an intelligent home usually applies sensors to sense the signals emitted from the human body, while video-based activity and fall recognition try to apply image processing and recognition algorithms on the recorded video or image data, such as [12].

In internal sensing, the devices are attached to the user, which leads to the research topic of signal-based activity recognition using wearable sensors, such as smartwatches and smartphones.
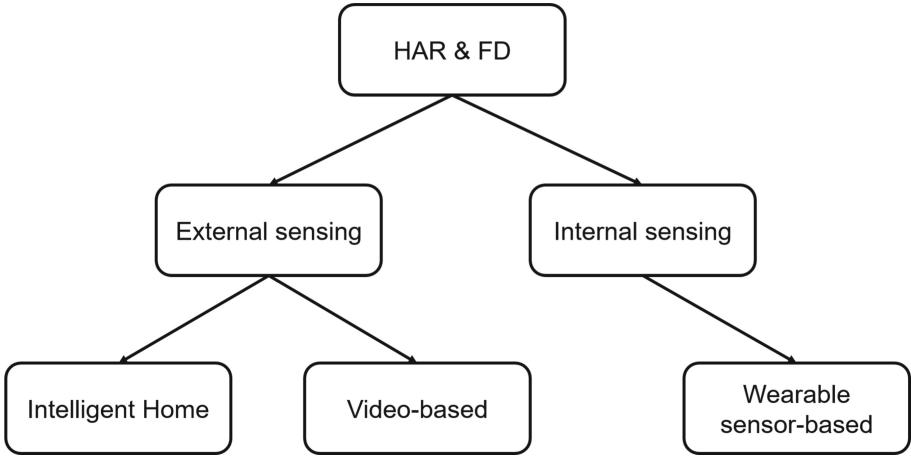
Fig. 1. Categories of sensing technologies for HAR and FD.

2   Hidden Markov Model and its Extension

2.1   Hidden Markov Model (HMM)

We denote T as the length of the observation sequence (i.e., $t = 1, 2, 3, ..., T$), the number of states S in the model by N and the number of observation symbols by M. A discrete observation HMM is then formally defined as 5-tuple $\lambda = (S, V, \pi, A, B)$ [9]:

 - $S = \{s_1, s_2, ..., s_N\}$: the set of all possible states;
 - $V = \{v_1, v_2, ..., v_M\}$: the discrete set of possible symbol observations;
 - $\pi = \{\pi(s_i)\}, \pi(s_i) = P(q_1 = s_i$ at $t = 1)$: the initial state distribution;
 - $A = \{a_{ij}\}, a_{ij} = P(q_{t+1} = s_j | q_t = s_i), 1 \le i, j \le N$: state transition probability distribution, usually represented as a matrix of probabilities, called transition matrix;
 - $B = \{b_i(k)\}, b_i(k) = P(V_k$ at $t|q_t = s_i), 1 \le i \le N, 1 \le k \le M$: observation symbol probability distribution, usually also represented as a matrix of probabilities, called emission matrix.

   There are three basic problems of interest that can be solved for the model to be useful in real-world applications [9]:

Problem 1: *Evaluation Problem.* Given the observation sequence $O = O_1, O_2, ..., O_T$ and a model λ, how do we efficiently compute $P(O|\lambda)$, the probability of the observation sequence, given the model?

Problem 2: *Decoding Problem.* Given the observation sequence $O = O_1, O_2..., O_T$, and the model λ, how do we choose a corresponding state sequence $Q = q_1, q_2, ..., q_T$ which is optimal in some meaningful sense (i.e., best "explains" the observation O)? HAR and FD are related to the solution of the second problem, for which a famous approach is the Viterbi algorithm [10].

Problem 3: *Optimization Problem.* How do we adjust the parameters A, B and π of the model λ to maximize $P(O|\lambda)$?

2.2   Continuous Density Hidden Markov Model (CDHMM)

In the HAR or FD task, the observations are the signals/segments/features instead of a finite set of discrete symbols. An obvious extension to the basic HMM model is to allow continuous observation space instead of a finite number of discrete symbols. In this model, parameters B (emission matrix) cannot be described as a simple matrix of point probabilities but rather as a complete Probability Distribution Function (PDF) over the continuous observation space for each state. Therefore, the values of $b_i(k)$ in the 5-tuple of HMM must be replaced with a continuous conditional probability distribution [13]:

$$b_j(x(t)) = P(x(t)|q_t = s_i), \forall x(t), i. \tag{1}$$

Equation 1 is called CDHMM. The conditional distributions (PDF) $b_j(x(t))$ can, theoretically, be arbitrary, but usually, they are restricted to be finite mixtures of simple parametric distributions, like Gaussians.

2.3   Hierarchical Hidden Markov Model (HHMM)

The HHMM applied state transition in a hierarchy-based structure. Thus, the definitions of states, observations, emission probabilities, and initial distribution in the 5-tuple of HMM or CDHMM remain while the transition pattern changes. In a traditional HMM, the transition reveals the relationship between two activities with probabilities, which could be regarded as a "linear" transition. In an HHMM, the entire structure is like a tree. In a high level of activity transition, some activities could have their low-level sub-activities. The sub-activities also have their transition relationship.

Taking the HMM modeling in [14] as an example, the "take a break" and "alternate workstation work" activities are at the same level. Thus, they have a transition relationship of probabilities with each other. Furthermore, the "take a break" activity could be sub-divided into several low-level sub-activities, such as "going on break" and "going off break", forming their individual transition topology. There is no direct path from the "going on break" activity to the "alternate workstation work" activity because they are neither in the same hierarchy nor in the same inheritance relationship.

3   HMM for Activity and Fall Modeling

3.1   State-of-the-Art Topology Study of HMM Human Activity and Fall modeling for Internal Sensing

HMM is powerful for modeling human activities, such as single motions like "walk," "stand," "sit," "jump," and "fall." Each activity can be modeled using a single HMM state [15,16], or a fixed number (greater than one) of HMM states [2,16,17], and both topologies work but with shortcomings.

Single-state. The single-state model for the "stand" activity may be enough, but for a more complicated motion like the "walk" and "fall" activities, the single-state may not be sufficient. The single-state modeling has also been applied successfully in a real-time HAR system [16].

Fixed Number of States. Some literature models each single motion activity with six [2] or ten [17] states, which achieves good recognition accuracy in different applications. However, these six or ten states are not interpretable, similar to neural network modeling. Moreover, there may exist redundancy (e.g., six HMM states of the "walk" activity may work well, but applying six HMM states to the "sit" activity should be over-engineering).

Distinct Number of States for Different Activities. Recently, state-of-the-art research tries to model different activities with different numbers and meanings of phases and states, called phase and state partitioning [11]. For example, one gait of the "walk" activity contains two phases and five states: the grounding-contact phase (three HMM sates) and the swing phase (two HMM states), as seen from one leg's view.

HMM State Generalization. Moreover, based on the above-introduced phase and state partitioning, some states can be shared among activities, like the situation in speech recognition. The generalized states are called Motion Units [11]. By applying this topology, falls can also be efficiently modeled with state generalization, and different types of falls [18] can be well recognized instead of only being detected as one activity, "fall."

3.2   Application of HMM for External Sensing

HMM also works for the external sensing, such as in an activity monitoring system for elderly care [19]. In this work, five types of wireless sensors, such as pressure mats, float sensors, and temperature sensors, are applied in the sensor network. The activities in this research are motion sequences rather than single motions, such as "leave," "shower," "sleep," and "drink." Four datasets are used to evaluate the activity recognition for elderly care. The HMM-based recognition results prove the capability of HMM modeling for external sensing of motion sequences, which helps assist elderly people.

3.3   HHMM in the Smart Home Application

HHMM, as introduced in Sect. 2.3, can also be well applied in an external sensing system for smart home technology. For example, the so-called MavHome project equips the rooms with several external sensors [14]. The system can control all lights, appliances, fans, heaters, and window blinds as an interactive result based on activity recognition called a data-driven approach for an intelligent environment. However, there are many challenges to solve for such an environment, such as sensor network failure, power outages, and slow and unreliable human behaviors.

4   Summary

This paper goes through the research framework and topics about HAR and FD, then briefly introduces HMM and its extension: CDHMM and HHMM, followed by the HMM modeling topologies for human activities in real applications with the description of several research works and projects.

References

1. Barandas, M., et al.: TSFEL: time series feature extraction library. SoftwareX **11**, 100456 (2020)
2. Hartmann, Y., Liu, H., Schultz, T.: Feature space reduction for multimodal human activity recognition. In: Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies, vol. 4, pp. 135–140. SciTePress, INSTICC (2020)
3. Hartmann, Y., Liu, H., Schultz, T.: Feature space reduction for human activity recognition based on multi-channel biosignals. In: Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies, pp. 215–222. SciTePress, INSTICC (2021)
4. Hopfield, J.J.: Artificial neural networks. IEEE Circ. Dev. Mag. **4**(5), 3–10 (1988)
5. Miikkulainen, R., et al.: Evolving deep neural networks. In: Computing Research Repository (CoRR), abs/1703.00548, arXiv (2017)
6. Yang, J., Nguyen, M.N., San, P.P., Li, X., Krishnaswamy, S.: Deep convolutional neural networks on multichannel time series for human activity recognition. In: International Joint Conference on Artificial Intelligence 2015, vol. 15, pp. 3995–4001. PMLR, Buenos Aires (2015)
7. He, K., Zhang, X., Ren,S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition 2016, pp. 770–778. IEEE, Las Vegas (2016)
8. Tuncer, T., Ertam, F., Dogan, S., Aydemir, E., Pławiak, P.: Ensemble residual network-based gender and activity recognition method with signals. J. Supercomput. **76**(3), 2119–2138 (2020)
9. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE **77**(2), 257–286 (1989)
10. Rabiner, L.R., Juang, B.H.: An Introduction to Hidden Markov Models. IEEE ASSP Mag. **3**(1), 4–16 (1986)

11. Liu, H., Hartmann, Y., Schultz, T.: Motion units: generalized sequence modeling of human activities for sensor-based activity recognition. In: 29th European Signal Processing Conference. IEEE, Dublin (2021)
12. Yang, L., Ren, Y., Zhang, W.: 3D depth image analysis for indoor fall detection of elderly people. Digital Commun. Netw. **2**(1), 24–34 (2016)
13. Continuous Density Hidden Markov Model. http://www.cs.helsinki.fi/u/ahonkela/dippa/node35.html/
14. Youngblood, G.M., Cook, D.J.: Data mining for hierarchical model creation. IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) **37**(4), 561–572 (2007)
15. Liu, H., Schultz, T.: ASK: a framework for data acquisition and activity recognition. In: Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies, vol. 3, pp. 262–268. SciTePress, INSTICC (2018)
16. Liu, H., Schultz, T.: A wearable real-time human activity recognition system using biosensors integrated into a knee bandage. In: Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies, vol. 1, pp. 47–55. SciTePress, INSTICC (2019)
17. Rebelo, D., Amma, C., Gamboa, H., Schultz, T.: Human activity recognition for an intelligent knee orthosis. In: 6th International Conference on Bio-inspired Systems and Signal Processing, pp. 368–371. SciTePress, INSTICC (2013)
18. Micucci, D., Mobilio, M., Napoletano, P.: UniMiB SHAR: a dataset for human activity recognition using acceleration data from smartphone. Appl. Sci. **7**(10), 1101 (2017)
19. Van Kasteren, T.L.M., Englebienne, G., Kröse, B.J.A.: An activity monitoring system for elderly care using generative and discriminative models. Pers. Ubiq. Comput. **14**(6), 489–498 (2010)