A study on a novel probabilistic procedure to efficiently encode progressive orders of activity patterns on Compact Features for Human Activity Recognition via Probabilistic First-Take-All

[1] M Ashish, [2]S. Prasad Babu Vagolu, and [3] Prof. Vedavathi Katneni
[1] PG Student, Department of Computer Science, GIS, GITAM,
[2] Asst Professor, Department of Computer Science, GIS, GITAM,
[3] Professor and HOD, Department of Computer Science, GIS, GITAM,

Abstract: With the attractiveness of mobile sensor technology, smart wearable devices open a unmatched location to solve the puzzling human activity recognition (HAR) problematic by learning communicative signs from the multi-dimensional daily sensor signals. This motivates us to improve a new algorithm relevant to both camera-based and wearable sensor-based HAR systems. Even though competitive classification correctness has been described, existing systems often face the encounter of distinguishing visually similar activities composed of activity patterns in different temporal orders. In this paper, we recommend a novel probabilistic procedure to efficiently encode progressive orders of activity patterns for HAR. Exactly, the algorithm studies an optimal set of latent patterns such that their progressive configurations surely matter in identifying different human activities. Then, a novel probabilistic First-Take-All (pFTA) methodology is announced to create solid features from the orders of these latent patterns to encode the entire sequence, and the temporal structural relationship between different sequences can be professionally dignified by the Hamming remoteness between compact features. Trials on three open HAR datasets display the suggested pFTA methodology can reach competitive presentation in relations of accuracy as well as efficiency.

***Index Terms* – Human activity recognition, temporal orders encoding, wearable sensors, learning to hash.**

I. Introduction

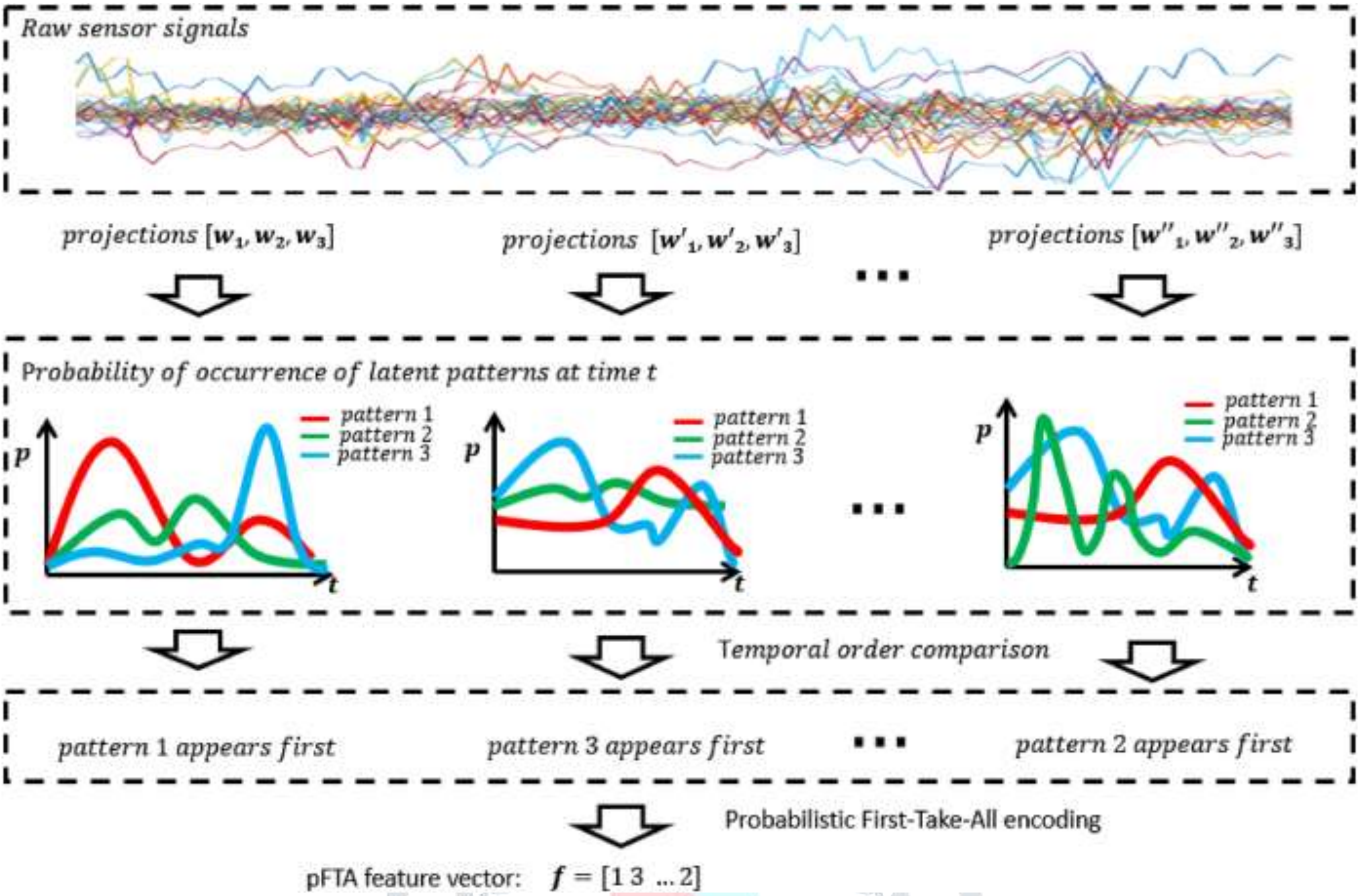
HUMAN Activity Recognition (HAR) has spawned intense researches in the past decade and continues to be an active research area. From the perspective of data acquisition, HAR can be classified into two categories: the camera-based systems using 2D or 3D cameras as the primary data capturing devices [1], [2], [3], and the wearable sensor-based HAR systems [4], [5], [6]. Compared with cameras-based HAR systems, mobile devices employ various miniature inertial sensors, such as accelerometers, gyroscopes, and magnetometers, which have advantages over cameras in availability, complexity, and privacy [7]. First, wearable sensors are body-worn and can provide sensing data virtually anytime and anywhere, while cameras may not have full coverage and the recorded videos are often subject to occlusions or visibility issues. Second, compared with heavy video data, signals from wearable sensors are lightweight, making them feasible for real-time online human activity detection and recognition on a large scale. Third, sensor data do not reveal user’s identity and are less vulnerable to privacy issues than the camera-based systems.

Considering the above merits, sensor-based HAR has attracted a great deal of interests and become an emerging research area with many important applications. For example, in the healthcare domain, wearable sensors can observe daily human activities, including sitting, standing, lying down, climbing floors and so forth, and have successfully detected abnormal activities in case of incidences [8]. With the development and increasing popularity of mobile technology, smart wristbands and smart phones provide a convenient solution to keep track of people’s daily and fitness activities such as counting steps and monitoring heart rate. Some smart wristbands even support auto-logging of different cycling events,1 or track the number of strokes in swimming.

In spite of many successful products and applications, sensor-based HAR remains a challenging problem due to its highly noisy nature and the complex temporal dynamics within the sequential sensor data. This makes it quite different from camera-based HAR where visual clues play an important role in capturing and recognizing human activities. Most existing sensor-based HAR approaches focus on extracting handcrafted statistical and structural features from the time domain and frequency domain of the multi-dimensional signals, such as mean, standard deviation, inter quartile rante (IQR), kurtosis, correlation, entropy and energy [5], [6], [9], [10], Fourier Transform coefficients [4], [5] and Discrete Cosine Transform coefficients. Although impressive accuracy has been reported, most of these features do not consider the temporal dynamics within the sequential data and therefore suffer from the confusion between activities with similar statistical patterns but different internal temporal structures

II. BACKGROUND

We will show that the pFTA scheme seeks to distil the most discriminative temporal information from the complete orders between latent patterns. This not only makes the pFTA features more compact than completely encoding the pattern orders, but also can improve its recognition accuracy by discarding the irrelevant pattern orders which are useless or may even play misleading roles in HAR. The compactness of the pFTA feature makes it extremely efficient for HAR using the nearest neighbour search based on the Hamming distance.



The main contributions of the paper are summarized as follows:

1) We introduce a novel probabilistic First-Take-All algorithm to exploit the underlying temporal dynamics in a human activity sequence from wearable sensor signals. A compact pFTA feature vector is produced from the above encoding algorithm, which carries discriminative power for human activity classification.

2) We further propose an optimization algorithm which can increase the discriminative capability of the pFTA feature by learning the most salient linear projections.

3) Extensive performance studies on three public HAR datasets show the proposed pFTA feature can outperform state-of-the-art statistical features on both accuracy and efficiency. pFTA can achieve comparative or even higher classification accuracy than deep recurrent neural networks such as LSTM with significantly higher efficiency.

III. EXISTING SYSTEM

The main contributions of the paper are summarized as follows:

1) We introduce a novel probabilistic First-Take-All algorithm to exploit the underlying temporal dynamics in a human activity sequence from wearable sensor signals. A compact pFTA feature vector is produced from the above encoding algorithm, which carries discriminative power for human activity classification.

2) We further propose an optimization algorithm which can increase the discriminative capability of the pFTA feature by learning the most salient linear projections.

3) Extensive performance studies on three public HAR datasets show the proposed pFTA feature can outperform state-of-the-art statistical features on both accuracy and efficiency. pFTA can achieve comparative or even higher classification accuracy than deep recurrent neural networks such as LSTM with significantly higher efficiency

- In spite of many successful products and applications, sensor-based HAR remains a challenging problem due to its highly noisy nature and the complex temporal dynamics within the sequential sensor data.
- It is worth noting that although we focus on sensor data from wearable devices in this paper, the proposed technique can potentially be used as a generic solution to other recognition problems on time series that contain rich temporal dynamics.
- We use the pairwise classification error in the training objective as a soft relaxation version of the inference so that we don't have to directly deal with a discrete optimization problem.

IV. PROPOSED SYSTEM

Extensive performance studies on three public HAR datasets show the proposed pFTA feature can outperform state-of-the-art statistical features on both accuracy and efficiency. PFTA can achieve comparative or even higher classification accuracy than deep recurrent neural networks such as LSTM with significantly higher efficiency. It is worth noting that although we focus on sensor data from wearable devices in this paper, the proposed technique can potentially be used as a generic solution to other recognition problems on time series that contain rich temporal dynamics. More recently, LSTM has been used in Human action recognition. For example, differential Recurrent Neural Networks (dRNN) has been proposed to model the dynamics of human actions by computing different-orders of derivative of state that are sensitive to the spatio-temporal structure of input sequence

- Compared with cameras-based HAR systems, mobile devices employ various miniature inertial sensors, such as accelerometers, gyroscopes, and magnetometers, which have advantages over cameras in availability, complexity, and privacy.
- To take advantage of the dynamics underlying human activities, one can resort to the Dynamic Time Warping (DTW) approach, which seeks to establish explicit correspondence between different sequences to align activity patterns in their temporal orders.
- Compared with heavy video data, signals from wearable sensors are lightweight, making them feasible for real-time online human activity detection and recognition on a large scale.
- Sensor data do not reveal user's identity and are less vulnerable to privacy issues than the camera-based systems.

V. EVALUATION

The procedure of the experiment and the evaluation results obtained are discussed in this section.

HUMAN ACTIVITY RECOGNITION

HUMAN Activity Recognition (HAR) has spawned intense researches in the past decade and continues to be an active research area. From the perspective of data acquisition, HAR can be classified into two categories: the camera-based systems using 2D or 3D cameras as the primary data capturing devices, and the wearable sensor-based HAR systems. Compared with cameras-based HAR systems, mobile devices employ various miniature inertial sensors, such as accelerometers, gyroscopes, and magnetometers, which have advantages over cameras in availability, complexity, and privacy. First, wearable sensors are body-worn and can provide sensing data virtually anytime and anywhere, while cameras may not have full coverage and the recorded videos are often subject to occlusions or visibility issues. Second, compared with heavy video data, signals from wearable sensors are lightweight, making them feasible for real-time online human activity detection and recognition on a large scale. Third, sensor data do not reveal user's identity and are less vulnerable to privacy issues than the camera-based systems.

INTERQUARTILE RANTE

In spite of many successful products and applications, sensor-based HAR remains a challenging problem due to its highly noisy nature and the complex temporal dynamics within the sequential sensor data. This makes it quite different from camera-based HAR where visual clues play an important role in capturing and recognizing human activities. Most existing sensor-based HAR approaches focus on extracting handcrafted statistical and structural features from the time domain and frequency domain of the multi-dimensional signals, such as mean, standard deviation, inter quartile rante (IQR), kurtosis, correlation, entropy and energy, Fourier Transform coefficients and Discrete Cosine Transform coefficients. Although impressive accuracy has been reported, most of these features do not consider the temporal dynamics within the sequential data and therefore suffer from the confusion between activities with similar statistical patterns but different internal temporal structures.

DYNAMIC TIME WARPING

To take advantage of the dynamics underlying human activities, one can resort to the Dynamic Time Warping (DTW) approach, which seeks to establish explicit correspondence between different sequences to align activity patterns in their temporal orders. Then human activities can be classified by grouping the similar sequences based on their alignments. However, it is usually computationally expensive to perform the DTW alignment, making it un scalable to a large dataset. Recently, Recurrent Neural Networks (RNNs), specially Long Short-Term Memory (LSTM) machines, become popular in modelling the temporal contexts of sequential data. Compared with handcrafted features, the LSTM models can learn expressive dynamic features from sequences in their memory states, and use these states to recognize different human activities.

FIRST TAKE ALL

This may limit their ability to fully explore the dynamic patterns whose temporal orders really matter in distinguishing different types of human activities. To address these challenges, we propose a novel probabilistic First-Take-All (pFTA) algorithm that explicitly explores the temporal dynamics underlying the sequential data and produces a highly compact feature representation from raw multi-dimensional signals. A diagram of the algorithm is illustrated. A set of linear projections are used to map sensor signals into different subspaces, each representing a latent pattern of human activity. Since each pattern is latent whose occurrence time in a sequence is a uncertain variable, we use the probability to characterize the moments that the pattern occurs in the sequence. We assume that these latent patterns are basic building blocks, which are combined to form ordered sequences to represent different types of human activities.

HYPOTHESIS AND PROBABILITY

Note that sequential data of the same category share similar temporal characteristics. We wish to extract and represent such unique traits by exploiting the aforementioned latent patterns in them. Individual latent patterns may appear at any time in the sequential data. However, the temporal orders between a set of K latent patterns in sequences of the same category are very likely to be consistent. For example, shows three patterns extracted from a 45-point data sequence. We can see that Pattern has high confidence scores on the first half of the sequence while pattern and pattern have high confidence scores on the second half of the sequence. It is this temporal order information that we want to encode to represent the temporal dynamics of the sequential data. Following this idea, we introduce the pFTA algorithm to extract the temporal dynamics by encoding the temporal orders of the latent patterns in the sequential data.

RESULTS

Paste screenshots

CONCLUSION

In this paper, we have addressed the challenge of HAR from the perspective of temporal order encoding and exploited the unique temporal dynamics underlying the sequential data. Extensive performance evaluations on two wearable sensor-based HAR datasets have demonstrated that the proposed pFTA feature has outperformed start-of-the-art statistical features in both accuracy and efficiency. In order to investigate the feasibility of pFTA as a generic solution to sequential data classification problems, we have also evaluated the proposed technique in a public RGBD human action recognition dataset. Results have shown that pFTA can achieve comparable results with respect to state-of-the art camera-based methods with higher efficiency

REFERENCES

[1] P. C. Ribeiro and J. Santos-Victor, "Human activity recognition from video: Modeling, feature selection and classification architecture," in Proc. Int. Workshop Human Activity Recognit.Modelling, 2005, pp. 61–78.

[2] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio,A. Blake, M. Cook, and R. Moore, "Real-time human pose recognition in parts from single depth images," Commun. ACM, vol. 56,no. 1, pp. 116–124, 2013.

[3] X. Yang and Y. Tian, "Super normal vector for activity recognition using depth sequences," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2014, pp. 804–811.

[4] K. Altun, B. Barshan, and O. Tunc¸el, "Comparative study on classifying human activities with miniature inertial and magnetic sensors," Pattern Recognit., vol. 43, no. 10, pp. 3605–3620,2010.

[5] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in Proc. Eur. Symp. Artif. Neural Netw. Comput.Intell. Machine Learn., 2013, pp. 437–442.

[6] J.-L. Reyes-Ortiz, L. Oneto, A. Sama, X. Parra, and D. Anguita,"Transition-aware human activity recognition using smartphones," Neurocomput., vol. 171, pp. 754–767, 2016.

[7] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," IEEE Commun. Surveys Tuts., vol. 15, no. 3, pp. 1192–1209, Jul.–Sep. 2013.

[8] J. Yin, Q. Yang, and J. J. Pan, "Sensor-based abnormal humanactivity detection," IEEE Trans. Knowl. Data Eng., vol. 20, no. 8, pp. 1082–1090, Aug. 2008.

[9] J.-Y. Yang, J.-S. Wang, and Y.-P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," Pattern Recognit. Lett., vol. 29, no. 16, pp. 2213–2220, 2008.

[10] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in Proc. Conf. Innovative Appl. Artif. Intell., 2005, pp. 1541–1546.

[11] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in Proc. 3rd Int. Conf. Knowl. Discovery Data Mining, 1994, pp. 359–370.

[12] S. Hochreiter and J. Schmidhuber, "Long short-term memory,"NeuralComput., vol. 9, no. 8, pp. 1735–1780, 1997.

[13] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," IEEE Trans. Inf. Technol. Biomed., vol. 10, no. 1,pp. 119–128, Jan. 2006.

[14] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher,"Activity recognition and monitoring using multiple sensors on different body positions," in Proc. Int. Workshop Wearable Implantable Body Sensor Netw., 2006, Art. no. 4.

[15] M. Ermes, J. Parkka, and L. Cluitmans, "Advancing from offline to online activity recognition with wearable sensors," in Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2008, pp. 4451–4454.

[16] O. Politi, I. Mporas, and V. Megalooikonomou, "Human motion detection in daily activity tasks using wearable sensors," in Proc. 22nd Eur. Signal Process. Conf., 2014, pp. 2315–2319.

[17] H. Wang, A. Kl€aser, C. Schmid, and C.-L. Liu, "Action recognition by dense trajectories," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.,
2011, pp. 3169–3176.

[18] J. Tang, X.-S. Hua, G.-J. Qi, and X. Wu, "Typicality ranking via semi-supervised multiple-instance learning," in Proc. 15th ACM Int. Conf. Multimedia, 2007, pp. 297–300.

[19] M. Baccouche, F. Mamalet, C. Wolf, C. Garcia, and A. Baskurt, "Sequential deep learning for human action recognition," in Proc. Int. Workshop Human Behavior Understanding, 2011, pp. 29–39.

[20] V. Veeriah, N. Zhuang, and G.-J. Qi, "Differential recurrent neural networks for action recognition," in Proc. IEEE Int. Conf. Comput. Vis., 2015, pp. 4041–4049.

[21] Y. Du, W. Wang, and L. Wang, "Hierarchical recurrent neural network for skeleton based action recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1110–1118.

[22] H. Hu, Z. Wang, J.-Y. Lee, Z. Lin, and G.-J. Qi, "Temporal domain neural encoder for video representation learning," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2017, pp. 2192–2199.

[23] H. Hu and G.-J. Qi, "State-frequency memory recurrent neural networks," in Proc. Int. Conf. Mach. Learn., 2017, pp. 1568–1577.

[24] J. Ye, H. Hu, K. Li, G.-J. Qi, and K. A. Hua, "First-take-all:Temporal order-preserving hashing for 3D action videos," CoRR, vol. abs/1506.02184, 2015.

[25] H. Hu, J. Velez-Ginorio, and G.-J. Qi, "Temporal order-based firsttake-all hashing for fast attention-deficit-hyperactive-disorder detection," in Proc. ACM SIGKDD Conf. Knowl. Discovery Data Mining, 2016, pp. 905–914.