Activity Recognition Based on Artificial Neural Network Approach using PIQ Robot

Bezobrazov S.,
Golovko V.
Brest State Technical University,
Brest, Belarus
Email: bescase@gmail.com
Email: gva@bstu.by

Kislyuk S.,
Sheleh A.
Octonion Technology,
Minsk, Belarus
Email: skislyuk@octonion.com
Email: asheleh@octonion.com

***Abstract*—This paper presents and explains an implementation of Learning Vector Quantization neural network for tennis activity detection and recognition using PIQ ROBOT device. The gesture recognition market is estimated to grow from 2013 till 2018 and is expected to cross $15.02 billion by the end of these five years. Analysts forecast the Global Gesture Recognition market to grow of 29.2 percent over the period 2013-2018. In terms of industry it means that currently consumer electronics application contributes to more than 99 percents of the global gesture recognition market. This paper starts our ambitious research in the area of artificial neural networks for activity recognition.**

Keywords*—*activity/gesture recognition, PIQ Robot, Artificial Neural Networks, LVQ neural network, time series classification.

I. Introduction

The gesture recognition market is estimated to grow at a healthy Compound Annual Growth Rate (CAGR) from 2013 till 2018 and is expected to cross $15.02 billion by the end of these five years. Analysts forecast the Global Gesture Recognition market to grow at a CAGR of 29.2 percent over the period 2013-2018. In terms of industry it means that currently consumer electronics application contributes to more than 99 percents of the global gesture recognition market [1].

There are a lot of researches on gesture recognition in the modern literature that incorporate the use of different types of sensors and models for several real life and virtual applications. On the basis of data acquisition, systems can be divided into two types: Vision-Based and Sensor-Based systems.

Vision-based recognition systems like Microsoft's Kinect [2], Nintendo Wii Remote [3], Oculus Rift [4], etc. cover several application areas such as surveillance, detection, control and other analysis of captured motion data. The problems with vision-based systems are of their high computational cost as most of them incorporate usage of GPUs, the range of use is limited by the camera viewing volume, a large number of cameras require to cover large spaces and light conditions and camera settings.

Glove-Based gesture recognition systems use a specific part or parts of the body to recognize gestures using a limited number of sensors. The striking example of such system is 'Acceleglove' [5] created by George Washington University. The systems based on the limited number of sensors have become more apparent in recent work because of the increasing popularity of wearable fitness devices. Unfortunately, such systems require a single training sample for each pattern and require users to define their own personal gestures.

State-of-the-art. The Hidden Markov Model (HMM) is the popular approach for gesture recognition. Thus Yamata et al [6] used discrete HMM and vector quantization for transforming the dynamic feature vectors to symbolic. Marcel et al. [7] trains an Input–Output HMM and apply it to recognize hand silhouette gestures.

The Finite State Machine (FSM) is another popular model in the area of gesture recognition. For example, Davis and Shah [8] proposed to use a FSM for modeling of different phases of a gesture.

Besides these models other approaches were implemented for gesture recognition. Thus, Suk et al. [9] propose to use dynamic Bayesian Networks to represent the relationship among gesture features based on motion tracking, and Flórez et al. [10] propose to use a self-organizing neural network to determine hand postures and gestures.

In this paper we present the first results of the implementation of artificial neural network approach for tennis sport activity detection and recognition. The structure of presented neural network can be implemented within the context of restricted computational resources. The rest of the paper is organised as follows. Section 2 gives the basic information about the characteristics of PIQ ROBOT as well as describes the artificial neural networks approach. Section 3 provides some experimental results for activity recognition for such kind of sport as tennis. The final section concludes this paper.

II. Artificial Neural Network approach for activity detection and classification

A. PIQ Robot

PIQ Robot is the ultimate sports tracker [11] (figure 1). PIQ integrates the latest inertial sensor technologies, BLE, NFC, pressure sensor and cutting edge microprocessor. PIQ Robot has been designed to be positioned in the ideal place for data collection whatever the sport. PIQ is a small, ultra-lightweight, waterproof and flexible multi-sport motion sensor connected to your smart device via BLE.

Figure 1. PIQ Robot

In scope of current research, PIQ Robot was switched to special engineering mode for having of access to inertial sensors data through BLE interface.

A special tool for Android smartphones, PIQ Log, was developed by Octonion Technology for reading of sensors data. It stores received time series into files (sensor data files) on memory of the smartphone.

Each sensor data file contains a header with service information. Sensor data files are friendly to comma-separated format and they can be easily exported to Microsoft Excel or any other tools for processing of CSV files.

Sensor data files contains the following fields:

1) Acceleration with gravity along 3 axis in local reference frame;
2) Angular velocity along 3 axis in local reference frame;
3) Quaternion for conversion of PIQ local coordinates to global coordinates.

PIQ Robot was fixed on wrist of tennis players during the game and transferred sensor data to PIQ Log application.

Sensors data associated with tennis strokes was used as a source information for training of neural networks at the following stage.

B. The structure of an ANN

Learning Vector Quantization (LVQ) neural network was selected for the classification task [12]. Since the internal resources of the PIQ Robot is quite limited (primarily because of the random access memory space) we have no possibility to use big-sized architectures of neural networks. In this case LVQ neural network provides good results using limited computational resources. The proposed architecture of neural network for activity recognition consists of three layers (Figure 2).

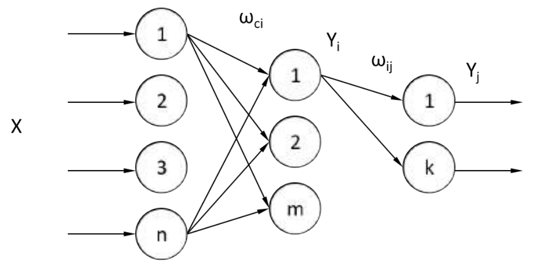

Figure 2. The structure of the ANN

The first (input) linear layer performs the function of distribution of an input signal to hidden layer. The number n of nodes in the first layer is defined by the dimension of the sliding window. The optimal dimension of sliding window directly depends on the nature of data and can be defined experimentally. In our case we use two different dimensions for sliding window. For such movements as Forehands, Backhands and Serves we use the dimension equal to 46. For Volleys activity the dimension of sliding window equals to 25. These parameters were chosen by the different length of different activities from collected sensor data files. Thus Serves are more long-term actions then Volleys, but Forehands and Backhands are almost equal and not so much differ from Serves.

The second layer (hidden layer) is competitive layer. It consists of m Kohonen neurons [12] and represents a vector quantization layer, which gives the cluster label of the input pattern. The competitive learning rule (winner-takes-all) is used for training of hidden layer [12]. During the training process neurons of hidden layers compete with each other. As a result the winner neuron (with the maximum weight) that characterized the class of data is defined. For this purpose the Euclidean space between input and weight vectors can be used:

$$D_i=|X-\omega_i|=\sqrt{(X_1-\omega_{1i}^2)+(X_2-\omega_{2i}^2)+...+(X_c-\omega_{ci}^2)},$$

where ω_{ci} is weight between c-th neuron of the input layer and i-th neuron of the competitive layer;

$X=[X_1, X_2, ..., X_c]$ is input pattern.

During the training process the synaptic connections are increased for winner neuron and stayed unchangeable for other neurons. Thus, after the training at the presenting the input pattern the activity of winner neuron is accepted to *"1"* while other neurons of hidden layers are accepted to *"0"*. The hidden layer training algorithm can be represented as the following set of steps:

1) Weight coefficients ω_{ci} of hidden neurons Y_i are initialized by random.
2) Input pattern is distributed on the neural network and next parameters are calculated:
   a) Euclidean distance between input and weight vectors of neuron elements from the Kohonen layer is calculated (equation 1)
   b) The winner neuron with the index k is defined

$$D_k = min_j D_j.$$

c) The weight coefficients of the winner neuron are modified according to the next equations:

$$\omega_{ck}(t+1) = \omega_{ck}(t) + \gamma(X_c - \omega_{ck}(t)),$$

if activity of the output neuron is corresponded to the class of inputted data;

$$\omega_{ck}(t+1) = \omega_{ck}(t) - \gamma(X_c - \omega_{ck}(t)),$$

otherwise.

3) The training process is continued while achieving of the desired accuracy.

The optimal number of hidden nodes depends on the data and can be defined experimentally, but it cannot be less than the numbers of classes.

The third layer is output layer, it consists of the linear neurons and performs the mapping of clusters from hidden layer into the existing classes. The activity of output neuron (when his value equal *"1"*) is characterized one or another class. For example, for the first class the activity of the 1-st neuron of output layer will be equal *"1"* while the activity of other neurons will be equal *"0"*. The number of neurons in hidden layers equals to the number of existing classes.

The described approach has some advantages:

1) The model is trained significantly faster than other neural network techniques like Backpropagation.
2) It is able to summarise or reduce large datasets to a smaller number of codebook vectors suitable for classification.
3) Able to generalise features in the dataset providing a level of robustness.
4) Can approximate just about any classification problem as long as the attributes can be compared using a meaningful distance measure.
5) Not limited in the number of dimensions in the codebook vectors like nearest neighbour techniques.
6) Normalisation of input data is not required (normalisation may improve accuracy if attribute values vary greatly).
7) The generated model can be updated incrementally.

All these advantages are very critical in the case of limited power resources.

III. Experimental results

PIQ ROBOT provides a broad spectrum of measurable parameters that contain (but is not limited by that) accelerations, linear accelerations, gyros, quaternions, pressure etc. Figure 3 shows the time series data that can be received from the sensor.

Time	AccelerX	AccelerY	AccelerZ	LinAccX	LinAccY	LinAccelZ	GyroX	GyroY	GyroZ	QuatW	QuatX	QuatY	QuatZ
227	1.475332	-2.9315	9.129815	0.119868	0.335458	-0.01677	-4	-5.125	-6	-0.70416	0.072327	0.168579	-0.68585
327	1.475332	-2.96982	9.158555	0.10414	0.295096	0.013583	-4.25	-5.125	-6	-0.70453	0.071777	0.169067	-0.68549
427	1.475332	-3.05605	9.177715	0.089115	0.206326	0.034099	-4	-5.125	-5.875	-0.70483	0.071228	0.169495	-0.68506
527	1.437012	-3.19017	9.129815	0.033229	0.070129	-0.01186	-3.75	-5	-6	-0.7052	0.070618	0.170044	-0.68463
627	1.456172	-3.17101	9.053174	0.033998	0.086372	-0.0867	-2.75	-4.875	-6.375	-0.70557	0.069946	0.170593	-0.6842
727	1.465752	-3.16143	9.110655	0.027452	0.092392	-0.02796	-2.375	-4.5	-6.875	-0.70593	0.069336	0.171021	-0.68372
827	1.437012	-3.34345	9.110655	-0.01227	-0.08027	-0.02289	-2.25	-4.125	-7.375	-0.70624	0.069397	0.171753	-0.68323
927	1.427432	-3.35303	9.120235	-0.09048	-0.08295	-0.00947	-0.75	-3.75	-8	-0.7066	0.069458	0.172302	-0.68274

Figure 3. The set of parameters from PIQ ROBOT

As can be seen, at each instant of time we have the set of parameters that characterize the changing the position of the sensor in the space as well as the its speed. We can detect and classify any specialized action (movements, strikes, ets) by analyzing these parameters.

For the training process of the proposed neural networks we use a data set that already marked by experts. It means that we know the start and end points (the first column in Table 1) for different types of movements. In our tests we chose tennis sport data. In tennis there are 4 types of basic strikes. Each type can contain several sub-types (see figure 4). As a result, we have 9 classes of different strikes in tennis that a grouped into 4 types.

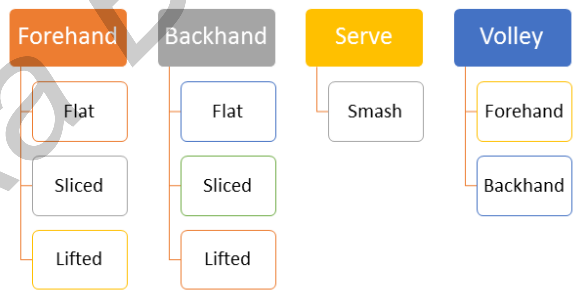

Figure 4. Existing tennis motions

For training of a neural network we used dataset "player_1" which contains marked data for different motions for one player. For the training of neural network we used 211 different motions. It is important to notice that different types of motion has different time period. Thus, Forehands motions take ≈ 185 ms on an average, Backhands ≈ 165 ms, Serves ≈ 230 ms and Volleys ≈ 124 ms. It means we care about the size of a sliding window and use different neural networks for different motion recognition.

In fact, in this work we use 4 neural networks. Here we can say about a combined classifier that consists of four neural networks. The first neural network is for detection and classification of Forehand, Backhand and Serve motions. The second neural network detects Volleys_forehand and Volleys_backhand motions. And finally, two neural networks classify the motions inside the Forehand and Backhand correspondingly. Thus we can detect and classify 9 possible different motions in tennis. Figure 5 represents the classification tree.

In the test for trained neural networks we used control files both for "player_1" and "player_2". Dataset

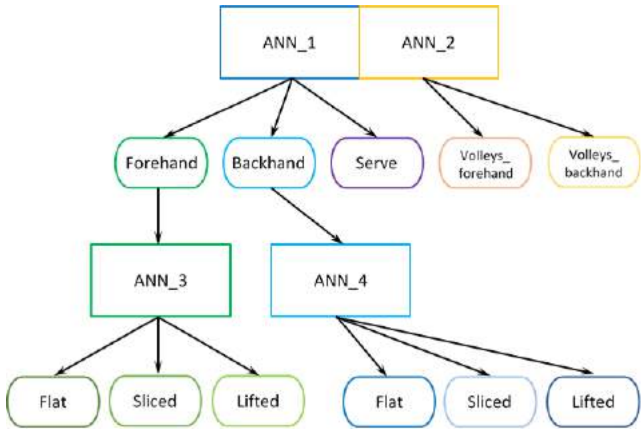

Figure 5. A combined classifier and classification tree

"player_1" contains activity of the same player that was "participated" in the training of neural network. Dataset "player_2" contains data from another player and his style of strikes can differs from player_1 significantly. The results of data classification are presented in table 1.

Table I. Classification results

Type of motion	Player_1, accuracy,%	Player_2, accuracy
Forehand	90	80
Forehand_flat	90	50
Forehand_sliced	90	90
Forehand_lifted	90	0
Backhand	90	90
Backhand_flat	90	0
Backhand_sliced	90	90
Backhand_lifted	90	0
Serve	90	90
Volley_forehand	90	90
Volley_backhand	90	90

It should be noted that on the current stage of research, the test files contain only a limited quantity of motions. As can be seen from the table 2, the proposed neural network shows good results for the main 4 motion types (Forehand, Backhand, Serve and Volley) equally good as for player_1, which data was used for training, and for player_2, which data absolutely new for classifier. For these motions we obtained the accuracy equals 90%. The accuracy for motions inside Volley (volley_backhand and volley_forehand) is also very good whether player_1 or player_2. But the results inside Forehand and Backhand types are not good even for player_1. In particular it can be explained by the similarity of flat and lifted motions and it is very difficult to distinguish them even if you do this manually.

IV. Conclusions

The paper presents the first results of an artificial neural network approach for activity recognition based on sensors data coming from PIQ Robot.

PIQ Robot is a powerful tool for mining and processing of inertial sensors data as it allows to apply different approaches to solving the problem of automatic segmentation and classification of sport motions.

In particular, this article demonstrates method of application of LVQ-nets to solving the problem of tennis motions classification. The implementation of neural networks for this task is quite new, innovative and carries a lot of advantages in comparison with traditional approaches. We showed that the combined classifier that contains several "small" Learning Vector Quantization neural networks can correctly detect and classify the basic motions using tennis dataset. Unfortunately, the proposed structure of neural network does not allow to solve the problem of detection of closely approximated motions like flat and lifted.

Based on the obtained results, it can be concluded that LVQ-nets are promising approach for enabling of automatic motions classification applied to sports.

References

[1] Need for Gesture Recognition – http://www.hcltech.com/engineering-rd-services

[2] Kinect – https://developer.microsoft.com/en-us/windows/kinect

[3] Wii remote – http://wii.com/

[4] Oculus Rift – https://www3.oculus.com/en-us/rift/

[5] Open-source data glove – https://www.technologyreview.com/s/414021/open-source-data-glove/

[6] J. Yamato, J. Ohya, K. Ishii, Recognizing human action in time-sequential images using hidden markov model, CVPR, 1992.

[7] S. Marcel, O. Bernier, D. Collobert, Hand gesture recognition using input–output hidden Markov models, FG, 2000.

[8] J . Davis, M. Shah, ECCV, Recognizing Hand Gestures, 1994.

[9] Suk, B. Sin, S. Lee, Recognizing hand gestures using dynamic bayesian network, FG, 2008.

[10] F. Flórez, J.M. García, J. García, A. Hernández, Hand gesture recognition following the dynamics of a topology-preserving network, FG, 2002.

[11] PIQ Robot - https://piq.com/

[12] Kohonen, T. Sell-organized formation of topologically correct feature maps / T. Kohonen // Biological Cybernetics. – 1982. – N43. – P. 59-69.

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЕ АКТИВНОСТЕЙ НА БАЗЕ PIQ ROBOT

Безобразов С., Головко В., Кислюк С., Шелег А.

Данная статья описывает нейросетевой подход для детектирования и классификации активностей в теннисе на базе устройства PIQ Robot. Прогнозируется, что в течение пяти лет рыночный спрос на технические средства для распознавания жестов будет расти и к 2018 году превысит сумму в 15 миллионов долларов. Статья описывает начало амбициозного проекта по использованию нейросетевого подхода для распознавания жестов и различного рода активностей используя сенсор PIQ Robot.