Recognition of boxing strike patterns using accelerometer data

Isaac Ambit Brao, Matej Razingar, Josep R. Zurita Nicolas
Escola Tècnica Superior d'Enginyeria Industrial de Barcelona, Barcelona Tech UPC, Barcelona, Spain
Submitted 10 June 2018

***Abstract*—This paper presents the complete process performed to achieve the identification of the multiple punch strikes commonly used in the sport of boxing. Acceleration data for each strike is obtained using the accelerometers smartphones are equipped with. The strikes being identified will be the straight punch, the hook and the uppercut, both left and right versions. To obtain all relevant acceleration data, two smartphones will be placed inside of the boxing gloves of the practitioners while they perform the exercises. The acceleration data will then be processed to extract singular instances of strikes, and a set of features will be extracted from each instance. These features will be used to train three different classifiers, which will be compared in their accuracy. Additionally, these classifiers will be tested with testing data specifically designed to resemble strike combinations that one may see in a boxing match.**

***Keywords*—Pattern recognition, classification, accelerometer, feature extraction, wavelet, movement recognition, support vector machine, random forest, k-nearest neighbors, boxing.**

I. Introduction

DURING the past ten years, cell phones have become an important part of daily life, and almost an apendice of ourselves. Some of the most important innovations on smartphones have been the incorporation of accelerometers and gyroscopes. Having such immediate access to the data from these sensors opens a new range of possibilities from athletes.

The objective of this project is to be able to differentiate among six classes of boxing strikes, obtaining the accelerometer lectures of two smartphones attached to the respective boxing gloves. The idea could be extended in a future to classify other kinds of strikes, movements or exercises, and to help professional athletes improve their technique. We follow a similar methodology to what has been previously used by others, in similar human movement classification problems.

We will first attach the accelerometers to the part of the body where the movements are most characteristic, then from the accelerometer data some features will be extracted and finally these features will be used to train multiple classifiers. The different features and classifiers will be evaluated separately with the purpose of finding the best combination. The different classes are shown in Figure 1 and are following:

1) Left Jab (or Left Cross if left-handed)
2) Right Cross (or Right Jab if left-handed)
3) Left Hook
4) Right Hook
5) Left Uppercut
6) Right Uppercut

The paper is organised in four sections:

- II describes the methods used to obtain the data, process it, extract the features and classify.
- III showcases the results.
- IV discusses the obtained results.
- V acknowledges the people and places that have helped with this project.

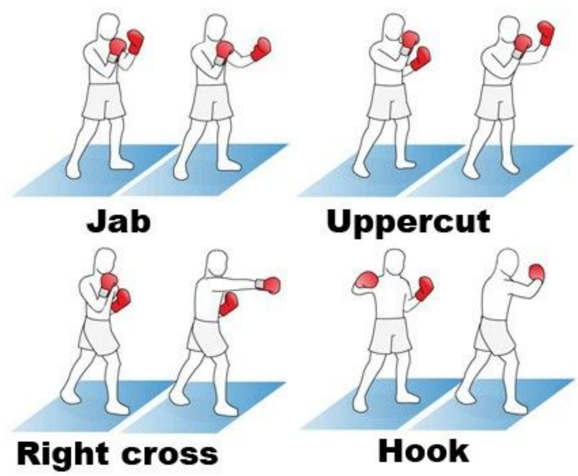

Fig. 1. Illustration of the different strikes.

II. Methods

The aim of the project was to classify the box punches obtained from two mobile phones. As the two accelerometers are separate devices, some acknowledgements had to be taken into consideration, and are presented in the following paragraph.

The mobile phones had to be positioned with the same orientation as features were extracted for each of the acceleration axes. Ideally, the phone should be placed to the same position on the testing subjects. The sampling frequency on both phones should be equal. The data had to be synchronized, as the aim was to classify the combination of punches.

The methodology consisted of collecting the data considering the previous acknowledgements, preprocessing the data, such as synchronizing, isolating and labeling the punches with associated class, feature extraction and classification.

A. Data collection

Accelerometer data was collected by two mobile phones, a 'Sony Xperia Z2' and an 'LG L70'. With these phones it is possible to sample accelerometer data with user-defined frequency. The accelerometer data was collected using the application 'Accelerometer Analyzer'. A sampling frequency of 50 Hz was selected for this study, as this was the maximum sampling frequency both mobile phones could reach. In previous studies, a frequency of around 50 Hz was used in order to classify daily activities, and as punching has a frequency similar to running, that was classified in previous studies, we assumed the sampling frequency used should be sufficient.

The accelerometer data consisted of three axes for each arm. In the study Preece et al. it was demonstrated that the optimal accelerometer placement was on the ankle. To interpret the activity classification problem to the punches problem, we supposed that the accelerometers should be placed near the point of impact. Thus, the phones were placed under the boxing glove on the wrist and fixed with the glove's strap, as it has been illustrated in Figure 2.

Fig. 2. Mobile phone fixed to the arm under the glove.

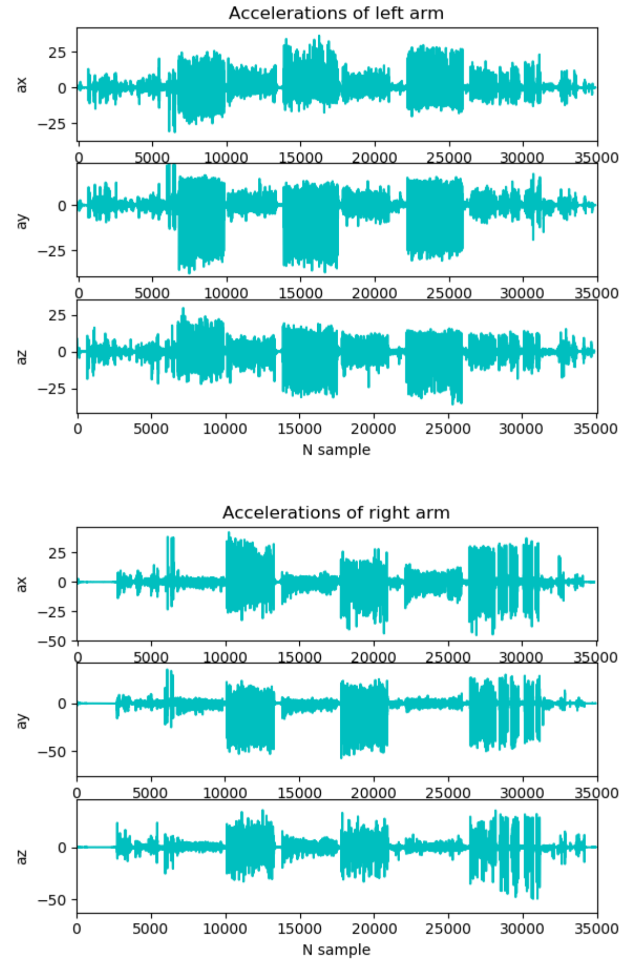

Fig. 3. Accelerations of left and right arm for one training set.

Three males and two female subjects participated in the study. The subjects had various experience in the boxing, two of them beginners and three intermediates. Three of the intermediate subjects completed testing set of about 50 punches per class while beginners completed about 25 punches per class. As there existed variability in number of punches among the classes the number of punches per class was limited to the recorded class with least punches, which equaled 191 punches per class. The testing data consisted of one set, performed by intermediate boxer. In the testing data some common boxing patterns were recorded. The patterns consisted of different combinations of two, three or four punches where each combination was repeated five times. The total number of punches in testing set was 180.

B. Preprocessing

The accelerometer data was in shape of two files, one for each arm and consisted of four columns, namely a_x, a_y, a_z and time from previous sample format. We observed that the data was not uniformly distributed as the time from previous sample ranged from 19 to 21 ms in contrast of expected 20 ms. Resampling to obtain uniformly distributed samples was performed. For some datasets, we also had to do synchronization as one phone started recording with a slight delay compared to another one. The example of training data format can be seen in Figure 3. This data is a complete training set of around 300 strikes. After the data was synchronized, the punches had to be isolated. The distinct acceleration peaks were observed on the y acceleration axis when a impact occurred. With this punch attribute we could place a window around the peak and isolate the punch with a certain window size. Various window sizes were tested. Our goal was to minimize the window size but still obtain solid results in testing. The window that seemed to suffice our conditions according to cross validation was 12 samples before and after the local peak, which results in a window of 0.48 s per punch using our sampling rate of 50 Hz. The example of extracting one punch from training data is shown in Figure 4.

With isolated punches we could label the training data to the according class. The testing data was captured and preprocessed in the same manner with exception of labeling it to the according class.

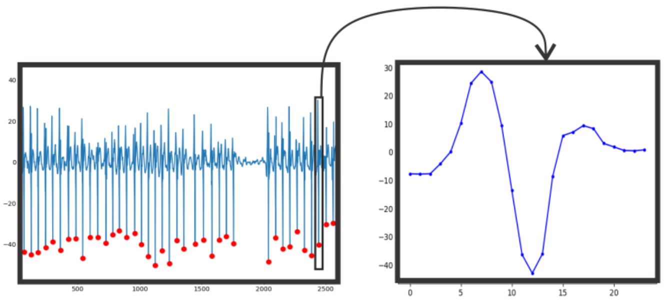

Fig. 4. Example of extracted punch from training data.

C. Feature extraction

A number of previous studies of activity classification based on accelerometer data derived the time domain features, frequency domain features and wavelet features.

Two groups of features were extracted. The first group contained time-domain features, namely mean, SD, median, minimum and maximum. Each of the features was extracted for all of the acceleration axes for both accelerometers (a total of 6 axes). In total, 30 time-domain features were obtained from the single punch.

The second group of features contained wavelet features, using discrete wavelet transform. Discrete wavelet transform decomposes the original time-domain signal initially into a coarse approximation and detail information by low-pass filtering (bandpass [0, f/2]) and high-pass filtering (bandpass [f/2, f]). In subsequent levels of decomposition, the approximation signal is split into a second approximation and detail coefficient. This process is repeated to the desired decomposition level [1],[3]. Most of the studies that were using discrete wavelet decomposition used wider window size than us, as it assures the decomposition to more detailed coefficients. However, we supposed that even less detailed wavelet coefficients would give us some information. Preece et al. also compared different wavelet features obtained from the wavelet coefficients. The best results were obtained using squared coefficients of wavelets and thus we used these wavelet features. Our window size provided us with five squared coefficients of wavelets for each acceleration axis, resulting in 30 wavelet features. The wavelet mother was determined experimentally, and the best results were obtained with the Daubechies 3 wavelet mother. All the features were scaled in order to center to the mean and to unit variance.

The results using only time-domain features, only wavelet features and both of them combined were obtained and compared as seen in the result section of this paper.

D. Classifiers used

In order to obtain as good results as possible several classifiers were compared, namely support vector machines, random forest and K-nearest neighbor classifiers. These classifiers were used because in previous articles of activity classification problems these classifiers were used. The support vector machine classifier was tested with radial basis function and linear kernels. The best results were obtained with linear kernel and parameter C = 1. K-nearest neighbor classifier was tested with various neighbors, the best results were obtained with K value of 3 neighbors. Random forest classifier was used with the default Python parameters. The three classifiers were compared as seen in Table I. The time that the machine required to classify the punches with different classifiers was similar for SVM and K-nearest, while Random Forest classifier was a bit slower, and also crashed in some experiments with certain punch window sizes. As it can be seen in Table I, the best performing classifier was K-nearest neighbor, thus being selected to classify the testing data.

TABLE I
COMPARISON OF GROUP OF FEATURES AND CLASSIFIERS USED

	Time Features	Wavelet features	Time + Wavelet
SVM	95 %	91 %	96 %
K-nearest	98 %	94 %	99 %
Random forest	95 %	88 %	96 %

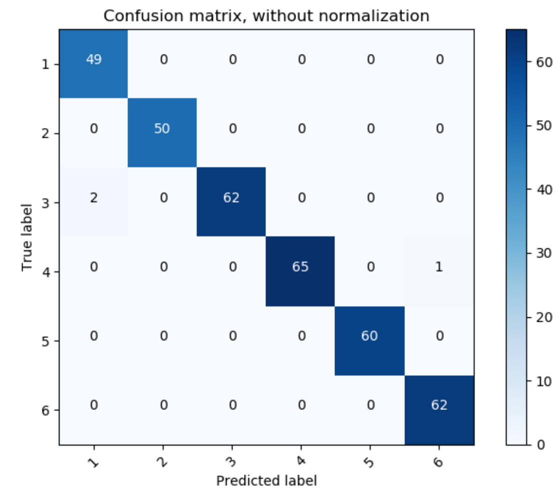

Fig. 5. Confusion matrix using 70-30 train-test split.

III. Results

Table I gives the classification accuracies using time domain features, wavelet features and both feature groups combined. The table also presents the difference in detection using different classifiers. The tests were done using cross-validation method, specifically, 10-fold cross-validation.

Overall, the highest classification accuracy (99% +/- 3%) was obtained when using both time and wavelet features and using K-nearest classifier with N=3 neighbors. Considering the sole group of features, the time domain features outperformed wavelet features. Furthermore, the difference between using both time and wavelet features was not significant compared to using just time domain features (98% +/- 3%). Figure 5 presents the confusion matrix using 70%-30% train-test split. Two left hook punches (class 3) were misclassified for left straight punch (class 1) and one right hook (class 4) was misclassified for right uppercut (class 6). It can be seen that the majority of punches were correctly classified. The testing data consisted of 180 punches done in combinations. The best

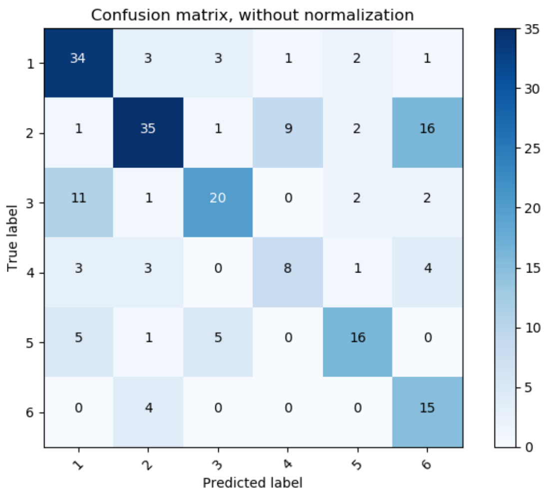

Fig. 6. Confusion matrix in testing data.

results were obtained when shortening the punch window to as low as 4 samples per punch. The accuracy of the testing set was 61 %. Figure 6 presents the confusion matrix for the testing data. We compared the detection between punching arms and obtained that punches are misclassified for another punch of the same arm in 64 cases and misclassified for punches of another arm in 17 cases. Table II presents the precision among the classes.

TABLE II
PRECISION AMONG THE CLASSES

Class	1	2	3	4	5	6	sum
Precision [%]	63	74	69	44	70	39	60

IV. Conclusions

In this paper we have presented the complete process followed to classify the most common boxing strikes. The results have been very positive, reaching a 99% of accuracy when performing cross validation on the training sets. This proves that the movements that we were aiming to classify are indeed very characteristic and can be easily identified. As [2] concluded, temporal domain features appear to be the most relevant, as an accuracy of 98% can be obtained from only the 5 time features we measured. In terms of the results of the testing sets, where we used the trained classifiers on realistic punch combinations, the results were significantly worse. The model performed better in determining the arm of punching than in the type of the punch, in other words the punch is more likely to be misclassified with another punch of the same arm than punch of another arm. We can assign this difference is accuracy not to an increased difficulty in identification of strikes but to a change in the conditions of the exercises. The classifiers were trained with individual strikes, resting between them. This allowed us to set a big window of time around the impact to gather features, but in the testing phase punches were being thrown at much higher speeds, the next strike commencing before the last one had ended. This meant we could not use the same time window and that some features could change. To obtain good results with realistic strike combinations, the classifiers should be trained with data recorded in these conditions, and if possible, with a greater sampling frequency.

V. Acknowledgements

We would like to thank Iron Gym Badalona and all the practitioners of boxing or kickboxing who regularly train there. But specially we would like to thank Ester, Elena, Javi, Moy, Nelson and Eduard for helping us gather the data that we used in this project.

References

[1] Preece SJ, Goulermas YJ, Member, IEEE, Kenney PJL, Howard D., *A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities from Accelerometer Data.*, IEEE Transactions on Biomedical Engineering 2009; 56:3.

[2] Ç.Berke Erdas, Isil Atasoy, Koray Açici, Hasan Ogul. *Integrating features for accelerometer-based activity recognition.*

[3] Christopher Torrence and Gilbert P. Compo, Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado. *A Practical Guide to Wavelet Analysis.*

[4] Ning Wang, Eliathamby Ambikairajah, Nigel H. Lovell, Senior Member, IEEE, and Branko G. Celler, Member, IEEE. *Accelerometry Based Classification of Walking Patterns Using Time-frequency Analysis,* in Proc. 29th Annu. Conf. IEEE Eng. Med. Biol. Soc., Lyon, France, 2007, pp. 4899-4902.