IMU-based Trick Classification in Skateboarding

Benjamin H. Groh
benjamin.groh@fau.de

Thomas Kautz
thomas.kautz@fau.de

Dominik Schuldhaus
dominik.schuldhaus@fau.de

Bjoern M. Eskofier
bjoern.eskofier@fau.de
Digital Sports Group, Pattern Recognition Lab
Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany

ABSTRACT

The popularity of skateboarding continuously grows for athletes performing the sport and for spectators following competitions. The presentation and the assessment of the athletes' performance can be supported by state-of-the-art motion analysis and pattern recognition methods. In this paper, we present a trick classification analysis based on motion data of inertial measurement units. Six tricks were performed by seven skateboarders. A trick event detection algorithm and four different classification methods were applied to the collected data. A sensitivity of the event detection of 94.2 % was achieved. The classification of correctly detected trick events provides an accuracy of 97.8 % for the best performing classifiers. The proposed algorithm holds the potential to be extended to a real-time application that could be used to make competitions fairer, to better present the assessment to spectators and to support the training of athletes.

Categories and Subject Descriptors

I.5.4 [**Pattern Recognition**]: Applications

Keywords

activity recognition, IMU, wearable sensors, sports

1. INTRODUCTION

The increasing interest in skateboarding as a competitive sport (e.g. by Street League Skateboarding [1]) requires new methods of motion analysis and innovative ways of presenting the performance of athletes. State-of-the-art signal processing technologies could offer an interdisciplinary development in skateboarding. The motion of the board can be captured by wearable sensors and analyzed in order to automatically detect and classify performed tricks. A trick classification system can be applied to competitions to support judges at the trick and scoring decision or to visualize the performance to the spectators. In addition, the system can find its application for non-professional skateboarders by providing feedback about their trick performance.

For the development of a trick-detection and classification system, sophisticated pattern recognition procedures are necessary. These methods are applied to the motion data which are often acquired by inertial measurement units (IMU). In contrast to video-based methods, IMUs do not require any external equipment and the acquisition is not confined to a specified area.

The first known approach to apply pattern recognition methods to skateboarding was proposed by Anlauff et al. [2]. They developed a real-time skateboarding game *Tilt 'n' Roll* for a Nokia N900 smartphone. No specific event detection algorithm for the real-time algorithm was implemented. Instead, all time stamps of the 70 Hz-sampled signal were considered as possible start of a trick and processed by the classification algorithm. The classification was based on linear discriminant analysis and was able to classify two skateboard tricks. An extension of the project was described by Reynell et al. [8]. They improved the hardware components of the system and visualized the skateboard movement based on accelerometer data. Further approaches for motion classification in the field of board sports were proposed considering several aspects. Harding et al. [6] developed a system to automatically calculate the number of rotations of halfpipe snowboarding tricks. Holleczek at al. [7] established a turn recognition for snowboarding. Based on specified gyroscope signal conditions, they distinguished between several snowboarding characteristics including the stance direction, turns and riding techniques. Sadi and Klukas [10] evaluated algorithms to detect jumps in snow sports including snowboarding. In a follow-up work, Sadi et al. [11] established a method to determine the air time of detected jumps.

All aforementioned approaches applied IMU signal processing methods to board sports with the objective to detect or classify specified activities. Most of them were established heuristically without applying machine learning algorithms. The only work containing machine learning algorithms in the field of skateboarding is the approach of Anlauff et al. It contained a two-trick classification based on one classifier without prior event detection. However, there was no comparison to other classifiers. In addition, the lack of an event detection could lead to unfeasible high computation times, especially for faster tricks that require a higher sampling rate than 70 Hz. In this work, we present a trick recognition pipeline containing an event detection and subsequent classification. Four different classifiers were applied to all trick events and compared for accuracy and computational effort.

KDD Workshop on Large-Scale Sports Analytics 2015, Sydney, Australia

2. METHODS

2.1 Data acquisition

2.1.1 Sensor hardware

The data for this study were collected with the sensor system miPod (Blank et al. [3]). The miPod system contains among others an inertial measurement unit (triaxial accelerometer and gyroscope). The accelerometer range was set to ± 16 g, the gyroscope was set to $\pm 2000\,°/\text{s}$. Measurements of both were obtained with a 16-bit resolution per axis. In addition, the internal real-time clock provided a timestamp with a synchronization accuracy of 150 ms. Data were collected with a sampling rate of 200 Hz. The data acquisition was video recorded with a Panasonic Lumix DMC-FT5 digital camera with a resolution of 640 x 480 pixels and a frame rate of 25 fps. The skateboards that were used for the data acquisition were provided by the skateboarders who participated in the study. One inertial sensor was attached behind the front axis of each board and adhered by 3M$^{\text{TM}}$ Dual Lock$^{\text{TM}}$ Reclosable Fasteners. By the attachment, the sensor-skateboard coordinate system was defined by the board's longitudinal axis x, the lateral axis y and the vertical axis z (Fig. 1).

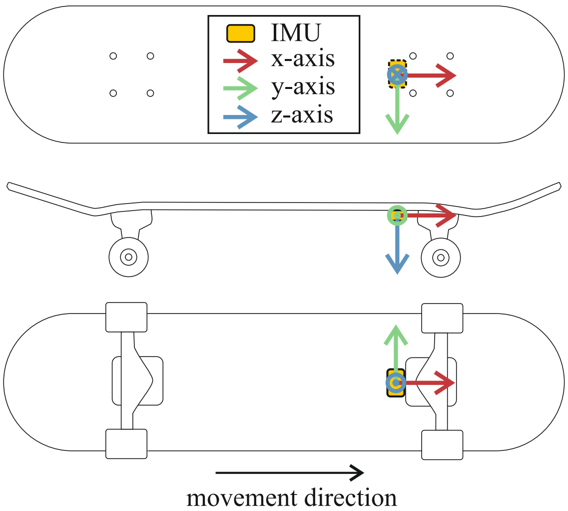

Figure 1: Sketch of the inertial sensor attachment including visualization of the sensor axes (not drawn to scale).

2.1.2 Study design

Seven experienced skateboarders (all male, age [years]: 25 ± 4, 3 *regular*, 4 *goofy*) participated in the data collection by performing six different tricks each. The given order of the trick execution was *Ollie*, *Nollie*, *Kickflip*, *Heelflip*, *Pop Shove-it* and *360-Flip* (Tab. 1). All tricks were repeated five times. If one trick execution was not performed correctly, the subjects could repeat the trick more than five times or move on to the next trick. The whole procedure was recorded on video camera as well as documented in a study protocol. In addition, the protocol contained the stance direction of the skater and the remark of an expert if a trick was performed correctly or not. A skateboarding glossary with an overview of relevant skateboard tricks and stance directions is provided in Tab. 1. The gyroscope signals of all performed tricks are visualized in Fig. 2. All subjects gave written consent to participate in the study and for the collected data to be published.

Table 1: Glossary with skateboard tricks and stance directions. The according trick signals are visualized in Fig. 2.

	Definition	Difference in *regular / goofy*
regular	stance direction front: left, back: right	–
goofy	stance direction front: right, back: left	–
Ollie (O)	nose up, tail up main rotation: y-axis ($+$y, $-$y)	none
Nollie (N)	tail up, nose up main rotation: y-axis ($-$y, $+$y)	none
Kickflip (K) / *Heelflip* (H)	360°-rotation main rotation: x-axis (kick: $-$x, heel: $+$x)	x-axis
Pop Shove-it (P)	180°-rotation main rotation: z-axis	x-axis z-axis
360-Flip (360)	*Kickflip* & 360° - *Pop Shove-it*	x-axis z-axis

2.2 Preprocessing

Depending on the stance direction, the obtained motion signal of some tricks varied. Comparing *regular* and *goofy* skaters, *Ollie* and *Nollie* did not change due to the main rotation about the y-axis. All the other tricks showed similarities in the y-axis but a mirrored behavior in the x- and z-axis (Tab. 1). In order to classify tricks of both types of skaters, the signals of the x- and z-axis of all goofy datasets were inverted in the preprocessing step.

2.3 Event detection

An event detection was implemented to determine relevant time intervals that included tricks. That approach was necessary to reduce the amount of data that were processed in the subsequent classification. For the event detection, the acceleration signal was segmented into windows. Based on considerations about the length of a trick and the duration of the landing impact, the length of the windows was set to 1 s with an overlap of 0.5 s.
The first step was to identify the landing after a trick. An event detection method was implemented to select windows that contained a possible landing impact. Therefore, the energy of the acceleration signal was calculated for each window as the sum of squares of all axes. A threshold for the detection of possible tricks was defined. If a window's energy exceeded the threshold, the window was selected for containing a possible landing impact and thus, for containing a possible trick. The threshold was determined by manually selecting and analyzing the first three performed *Ollies* of each skater. Three *Ollies* were chosen as threshold decision because its landing impact is rather small compared to other tricks and it can be performed by most skateboarders. The energy values of the windows that contained the three tricks were calculated and the lowest value was selected. The threshold was then set to a level of 10 % below the lowest

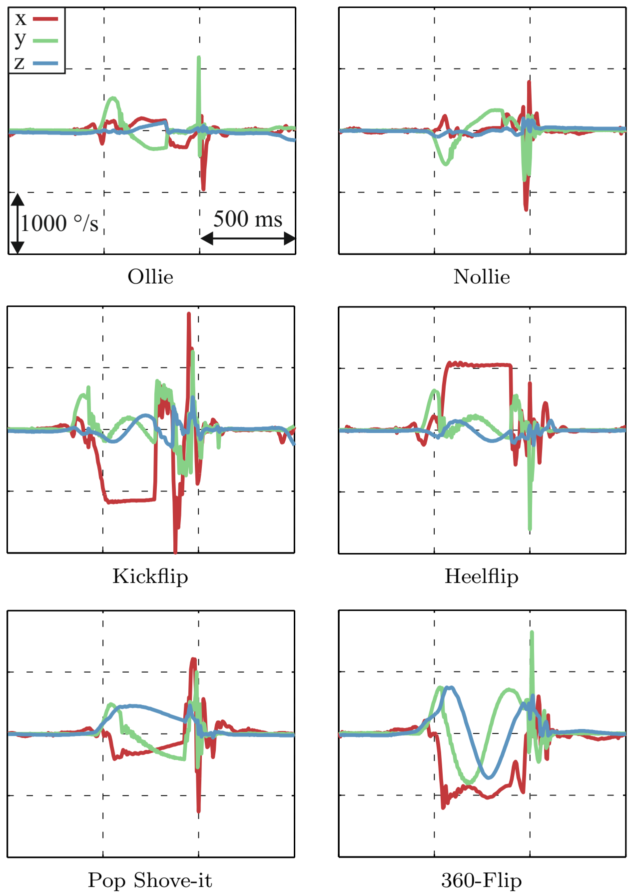

Figure 2: Example gyroscope signals of one subject of all tricks that were used for the classification analysis.

determined value.
In the next step, the actual trick interval was defined for each selected window. Therefore, the exact time step t_{land} of the landing impact was computed as the maximum of the acceleration signal energy within each selected window. From the landing time step t_{land} the trick interval was defined from $[t_{land} - 1\,\mathrm{s}]$ to $[t_{land} + 0.5\,\mathrm{s}]$ in order to include the preparation of the trick and the trick itself before the landing impact and some instances after the landing. The detected trick intervals were provided for the next processing step, the feature extraction.

2.4 Feature extraction

Feature vectors were calculated for all trick intervals. Statistical features included mean, variance, skewness and kurtosis. Frequency features contained the dominant frequency and the bandwidth. All features were calculated for the three acceleration axes, the three gyroscope axes and the norm of the acceleration and gyroscope signal. In addition, the x-y-correlation, the x-z-correlation and the y-z-correlation were calculated for each sensor type. Hence, 54 features were computed for each window in the input data.

2.5 Feature selection and classification

The feature selection and the classification were both performed by the Embedded Classification Software Toolbox (ECST) [9]. The ECST implementation is partially based on Weka [5] but additionally provides an analysis of required arithmetic operations in order to estimate the computational effort of each classification approach. From the 54 extracted features, only the best performing ones were used for the classification. Therefore, a feature selection was performed by ECST with a best-first forward selection. Four classifiers were compared: Naive Bayes (NB), Partial Decision Tree (PART), Support Vector Machine (SVM) with a radial basis kernel and k-nearest neighbor (kNN). For SVM and kNN, a grid search for the best performing parameters was executed. SVM was optimized for the parameters γ and C in the ranges $\gamma \in [2^{-5}; 2^{5}]$ and $C \in [2^{-5}; 2^{5}]$. kNN was analyzed for k $\in \{1,3,5\}$.

2.6 Evaluation

The goal of the trick event detection was to correctly detect trick events and to ignore non-trick time intervals. The proposed procedure was performed for all subjects. The resulting intervals were compared to manually labeled trick events in the video recording. For the evaluation, sensitivity and specificity for the detection of trick events were calculated in relation to the number of all segmented windows.
In this early stage of the project, the evaluation of the classification was only based on correctly detected trick events. Non-trick events that were incorrectly selected by the event detection algorithm were excluded from further processing. The trick events were classified by the four classifiers in order to obtain the accuracy and computational effort (computation time and required operations for one run without grid search) of each of them. The evaluation was based on a leave-one-subject-out cross-validation.

3. RESULTS

The total number of segmented windows for the event detection was 13542 containing 343 trick windows. The algorithm correctly detected 323 of the 343 events and incorrectly detected 13 events that did not contain a trick. This results in a sensitivity of 94.2 % and a specificity of 99.9 %. The classification was based on the 323 correctly detected tricks. The results of the ECST-based analysis of the accuracy and computational effort of all classifiers are summarized in Tab. 2. The best overall accuracy was achieved for Naive Bayes and SVM with 97.8%. The confusion matrix of the Naive Bayes classification is provided in Tab. 3.

Table 2: Overall accuracy and computational effort of all classifiers.

	NB	**PART**	**SVM**	**kNN**
accuracy [%]	97.8	93.4	97.8	96.0
computation	low	low	high	middle
- operations:	360	41	1015	1086
- time [s]:	6.2	10.6	32.7	5.2

4. DISCUSSION

The event detection results show a sensitivity of 94.2 % and a specificity of 99.9 %. However, it has to be considered, that the skateboard training was performed for the purpose of this study and lasted only about 15 minutes per subject. The subjects were asked to execute the tricks in a regular

Table 3: Confusion matrix of the trick classification with NB.

predicted \ true	O	N	K	H	P	360
O	33	1	0	0	0	0
N	2	31	0	0	0	0
K	0	0	36	0	0	1
H	0	0	0	37	0	0
P	0	0	0	0	35	1
360	0	0	0	0	0	49

manner but it can be assumed that a typical skateboarding training of one or two hours would contain more non-trick events than this study. This fact might lead to a decreased specificity in a real training scenario. However, a slightly lower specificity by detecting more non-trick events might not influence the system's final performance considerably.
The best classification results were achieved with Naives Bayes and SVM with an accuracy of 97.8 %. Classification errors that can be seen in the confusion matrix occurred for *Ollie* and *Nollie* and for two *360-Flips* that were mistakenly classified as *Kickflip* and *Pop Shove-it*. The confusion of *Ollie* and *Nollie* can be explained by their similar signals. The misclassification of the *360-Flips* could be a result of the *360-Flip* consisting of a combination of *Kickflip* and *Pop Shove-it* (Tab. 1 and Fig. 2). An extension of the extracted features by trick specific features (e.g. change of orientation) could solve these issues.
Despite the high classification rate, it has to be considered that all results were obtained by only processing actual trick events with correctly performed tricks. Non-trick events that were mistakenly detected by the event detection (false positives) were not considered for further processing. In the final application of the proposed system, a further method has to be found to overcome this limitation. In addition, not-correctly performed tricks have to be considered for the system. Their signal might result in similar features but can still contain small deviations. Possible solutions for the non-trick events could be the implementation of a null-class or a trick/non-trick decision based on the classification probability. A similar solution has to be found for the correctly and not-correctly performed tricks.
Considering a real-time trick classification application, the computational effort would have to be analyzed in detail [4]. This study already showed the relative behavior of computation time of different classifiers but an advanced computation time analysis in comparison to the classification accuracy would be necessary. In addition, the results were achieved by the classification of only six tricks. For an actual application, a higher variety of tricks would be required.

5. SUMMARY AND FUTURE WORK

We conducted a study with seven subjects performing six skateboarding tricks. The data were collected by accelerometer and gyroscope sensors. Possible trick events were determined by a window-based approach. All correctly detected trick events were classified by four classifiers. The best performance was obtained with a classification by Naive Bayes and Support Vector Machine with an accuracy of 97.8 %. With our study, we provide a reliable base for future work in further analyzing the performance of a trick recognition system in skateboarding or board sports in general. Further developments could include the implementation of the system in a real-time application and an extension of the application by a trick performance rating. Thereby, competitions could be supported by analyzing, classifying and rating the skateboard motion. Furthermore, amateur skateboarders could use a classification system as a training device and to exchange their achieved performance.

6. ACKNOWLEDGMENTS

The authors would like to thank *Bretterbude Skateshop*, Erlangen and all participants of the study for their support during the data collection.

7. REFERENCES

[1] *Street League Skateboarding*. [Online]. http://www.streetleague.com/ [Accessed: January 31, 2015].

[2] J. Anlauff et al. A method for outdoor skateboarding video games. In *7th Int. Conf. on Advances in Computer Entertainment Technology (ACE)*, pages 40–44, 2010.

[3] P. Blank, P. Kugler, H. Schlarb, and B. Eskofier. A wearable sensor system for sports and fitness applications. In *19th Annu. Congr. of the European College of Sport Science (ECSS)*, page 703, 2014.

[4] B. Eskofier, M. Oleson, C. DiBenedetto, and J. Hornegger. Embedded surface classification in digital sports. *Pattern Recognition Letters*, 30(16):1448–1456, 2009.

[5] M. Hall et al. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18, 2009.

[6] J. W. Harding, C. G. Mackintosh, A. G. Hahn, and D. A. James. Classification of aerial acrobatics in elite half-pipe snowboarding using body mounted inertial sensors (P237). *The Engineering of Sport*, 7(2):447–456, 2008.

[7] T. Holleczek, J. Schoch, B. Arnrich, and G. Troster. Recognizing turns and other snowboarding activities with a gyroscope. In *14th Int. Symp. on Wearable Computers (ISWC)*, pages 1–8, 2010.

[8] E. Reynell and H. Thinyane. Hardware and software for skateboard trick visualisation on a mobile phone. In *Conf. of the South African Institute for Computer Scientists and Information Technologists (SAICSIT)*, pages 253–261, 2012.

[9] M. Ring, U. Jensen, P. Kugler, and B. Eskofier. Software-based performance and complexity analysis for the design of embedded classification systems. In *21st Int. Conf. on Pattern Recognition (ICPR)*, pages 2266–2269, 2012.

[10] F. Sadi and R. Klukas. Reliable jump detection for snow sports with low-cost mems inertial sensors. *Sports Technology*, 4(1-2):88–105, 2011.

[11] F. Sadi, R. Klukas, and R. Hoskinson. Precise air time determination of athletic jumps with low-cost MEMS inertial sensors using multiple attribute decision making. *Sports Technology*, 6(2):63–77, 2013.