Wearable Trick Classification in Freestyle Snowboarding

Benjamin H. Groh, Martin Fleckenstein, Bjoern M. Eskofier

***Abstract*— Digital motion analysis in freestyle snowboarding requires a stable trick detection and accurate classification. Freestyle snowboarding contains several trick categories that all have to be recognized for an application in training sessions or competitions. While previous work already addressed the classification of specific tricks or turns, there is no known method that contains a full pipeline for detection and classification of tricks from multiple categories. In this paper, we suggest a classification pipeline containing the detection, categorization and classification of tricks of two major freestyle trick categories. We evaluated our algorithm based on data from two different acquisitions with a total number of eleven athletes and 275 trick events. Tricks of both categories were categorized with recall results of 96.6 % and 97.4 %. The classification of the tricks was evaluated to an accuracy of 90.3 % for the first and 93.3 % for the second category.**

I. Introduction

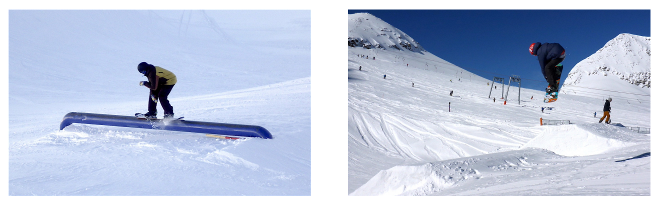
Digital motion analysis in snowboarding leads to a deeper insight in the sport, can enhance the performance of athletes and support the judges' decisions at competitions [1]. However, to enable an automated digital motion analysis, tricks first have to be detected reliably and classified accurately.

An unobtrusive and cheap method for these goals is based on the data processing of inertial-magnetic measurement units (IMMUs). These devices provide accelerometer, gyroscope and magnetometer measurements. Based on IMMUs attached to a snowboard, the current state of the board (rest, in motion, etc.) as well as the board's orientation can be monitored at all times. This information can be used for an automated trick detection and classification.

In literature, various approaches for snowboard trick analysis were published. The group of Harding et al. focused on trick analysis in half-pipe snowboarding [2], [3]. They calculated the rotation angles of the tricks during the air time and could distinguish between tricks with varying rotations. Furthermore, Holleczek et al. [4] proposed an algorithm to determine several snowboard turns and the current riding style by processing of GPS and gyroscope data. Similar studies were conducted in the related field of skateboarding. Anlauff et al. [5] proposed a two-trick classification system which was implemented to run in real-time. In addition, our group developed a pattern recognition-based algorithm for skateboard trick detection and classification in six different trick classes [6].

The aforementioned studies were always focused on only one trick category (e.g. only half-pipe tricks). However, freestyle snowboarding implicates the athletes' freedom of decision for tricks from several categories in one run. Hence, an extension to previous work is required that is capable of classifying tricks from multiple categories. Furthermore, all snowboarding-related approaches were based on decision making by single assumptions (e.g. rotation about one axis) and hence are hardly extendable to a wider variety of more advanced tricks.

In this work, we propose an algorithm that first detects possible trick events, then distinguishes between two of the main freestyle trick categories (grind tricks and air tricks, see Fig. 1) and finally classifies the actual trick. We implemented the algorithm based on pattern recognition methods, which are extendable in both the trick categories and the trick classes.



(a) Grind (b) Air

Fig. 1. Example of snowboard trick categories grind and air. Grinds are performed by sliding over static objects, e.g. rails or boxes. Airs are jumps using a kicker and performing defined rotations during the air time.

II. Methods

A. Data acquisition

1) Hardware: The data acquisition of this study was based on the *miPod* inertial-magnetic measurement unit (Blank et al. [7]). It contained an *InvenSense MPU-9150*, which provides sensing in three-axes by an accelerometer, a gyroscope and a magnetometer. The accelerometer range was set to $\pm$ 16 g, the gyroscope was set to $\pm$ 2000 °/s and the magnetometer measured in the range of $\pm$ 1200 μT. Data were obtained with a 16-bit resolution per axis. A *miPod* device was attached to each board behind the front binding. It was adhered by 3M^{TM} Dual Lock^{TM} Reclosable Fasteners and in addition fixed with duct tape. The sensor was attached in a way that its x-, y- and z-axes represent the longitudinal, the lateral and the vertical axis (see Fig. 2). In addition to the IMMU-device, two *GoPro Hero 2* cameras were used during the data collection. They were set to a resolution of 848 x 480 pixels with a frame rate of 50 Hz.

2) Study design: Data were collected in two different acquisitions. A short trick description and the number of executions in both acquisitions are summarized in Tab. I.

All authors are with the Digital Sports Group, Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
Contact Email: benjamin.groh@fau.de

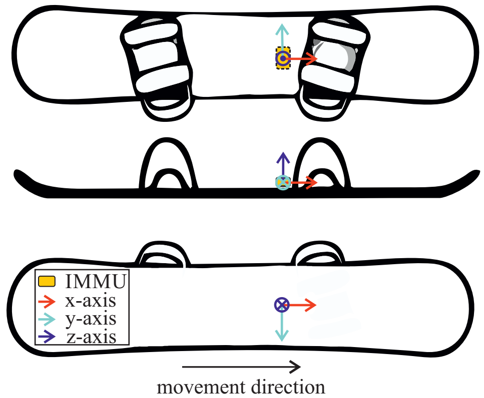


Fig. 2. Coordinate system of board-sensor-system.

TABLE I
OVERVIEW OF TRICK CATEGORIES AND CLASSES OF BOTH ACQUISITIONS. THE ROTATION AXES REFER TO REGULAR STANCE DIRECTION.
BS: BACKSIDE, *FS*: FRONTSIDE

acquisition	description / rotation	number
A: Hintertux		
Grinds (category I)	sliding on obstacle (e.g. box)	12
Airs (category II)	jump over kicker with air time	62
B: Bispingen		
Grinds (category I)		
- 50-50	straight slide, no rotation	38
- BS-Boardslide	90° (+z), slide, 90° (-z)	37
- FS-Boardslide	90° (-z), slide, 90° (+z)	32
Airs (category II)		
- Method	straight jump, no rotation	37
- BS-180	180° (-z)	32
- FS-360	360° (+z)	25

All participants were informed of snowboard-related risks and gave written consent to participate in the study and for the collected data to be published.
Acquisition A was performed in Hintertux, Austria. Four subjects (age in years: 18 ± 9, size in cm: 167 ± 13, all male, stance direction: two regular and two goofy) participated in the study by performing freestyle snowboard tricks. The study contained both grind and air tricks. In total, data of twelve grinds and 62 airs were collected. However, due to a large variety of different tricks of the athletes, there was no differentiation feasible between tricks of one category in this acquisition.
In contrast, acquisition B was conducted in order to additionally differentiate between tricks in both categories. Therefore, seven further athletes (age in years: 15 ± 8, size in cm: 156 ± 13, all male, stance direction: four regular and three goofy) performed three different grinds and three different airs. Data of approximately five runs per athlete were collected which leads to a total number of 201 performed tricks. Not all athletes performed the same number of runs nor did they always execute the same number of tricks in each run.
Both acquisitions contained a specified calibration procedure. For the accelerometer and gyroscope calibration, the IMMUs were set in six static positions and rotated about each of the three axes. A calibration of the magnetometer was not necessary. The magnetometer signal was only analyzed for relative changes and not for absolute measurements of single axes. All data of the trick performances and the calibration procedure were stored for later processing.

B. Preprocessing

The obtained data were preprocessed for two purposes: calibration and correction of the stance phase influence.
Based on the aforementioned calibration procedure, the accelerometer and gyroscope data were calibrated by using the static measurements and the rotations about all axes. The implemented calibration method followed the algorithm of Ferraris et al. [8].
Furthermore, a correction of the stance phase influence was necessary in order to obtain consistent measurements of all subjects. Snowboarders either ride regular or goofy which defines the leading foot during motion [9]. Comparing regular and goofy riding subjects, the obtained data did not vary for rotations about the y-axis but showed mirrored behavior in the x- and z-axes. In order to analyze tricks of both types, the signals of the x- and z-axes of all goofy data sets were inverted before further processing.

C. Event detection

An event detection was developed in order to extract possible trick events. These events were found by a distinct peak in the accelerometer signal produced by the landing impact after a trick. In order to overcome the influence of sensor noise, the L^1-norm (sum of the absolute values of all axes) $s_{a,t}$ was computed for all times t. Subsequently, the signal was analyzed by a window-based threshold-approach with a window length of 50 samples (0.25 s) and an overlap of 49 samples. Every time $s_{a,t}$ exceeded a pre-defined threshold ϵ, the corresponding signal window was considered for further processing. The threshold ϵ was determined by a leave-one-subject-out cross-validation (LOSO-CV) with the criterion of detecting all trick events of the subjects in the training data set of each LOSO-CV iteration.

D. Classification of trick category

Each detected event was processed in order to analyze if it contained a grind trick event (category I), an air trick event (category II) or an incorrectly detected no-trick event (category III). Grinds and airs showed typical features that could be used for the differentiation (see Fig. 3). Grinds were always performed on a metallic rail or a metallic box. Hence, the magnetometer signal showed a distinct behavior that varied from the normal run aside from grind objects. Airs always included an air time which could be noticed by a low noise level in the accelerometer signal.
The classification was performed in a two-stage approach. In the first step, grind events were classified with the magnetometer signal. In the second step, all events that were not classified as grind were further analyzed with the accelerometer signal for containing an air event. It was assumed that

the most relevant information for the category classification could be extracted from the time interval $\Delta t_{rel,cat}$ before a detected landing impact. Furthermore, there was no differentiation between single axes and only the L^1-norm of both the magnetometer signal $s_{m,t}$ and the accelerometer signal $s_{a,t}$ were used. A more detailed description of the two-stage approach is provided in the following paragraphs.

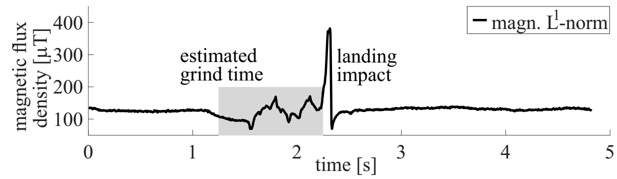
1) Grind classification: The grind classification was based on the high variance of the magnetometer signal influenced by the metallic surface. Therefore, the variance of the magnetometer L^1-norm $s_{m,t}$ was calculated with a moving-window approach with a window length of 10 samples (0.05 s). This window length was chosen in order to eliminate noise while sustaining relevant changes in the signal. The only feature that was extracted was the number of samples that exceeded a defined threshold $\eta_{var,mag}$ in the computed variance signal. The grind classification was implemented by a Naive Bayes classifier which distinguished between two classes: grind events (category I) and all other events (category II and III). The implemented classifier was based on an adapted cost sensitivity matrix [10] in order to weight the recall of the classification result over the precision. The corresponding proportion of cost$_{false_negative}$ to cost$_{false_positive}$ was defined by the weight factor $f_{w,grind}$.

2) Air classification: Due to the already classified grind events, some incorrectly detected events (category III) could be eliminated. Based on the assumption, that there is always a dead time between two performed tricks, all events, which were detected within the time interval of Δt_{pause} before and after a classified grind of the same athlete, were excluded from further processing. Subsequently, an air classification was performed similarly to the previous grind classification. The variance of the accelerometer L^1-norm was calculated with a moving-window approach of window length 10 samples (0.05 s). The only used feature was based on the low signal variance of air tricks during the air time. In order to detect samples during the air time, a threshold $\eta_{var,acc}$ was defined. The number of samples that were below this threshold was used as the extracted feature. Analogue to the previous approach, a Naive Bayes classifier was trained using a cost sensitivity matrix. The cost proportion was defined by the weight factor $f_{w,air}$, respectively.

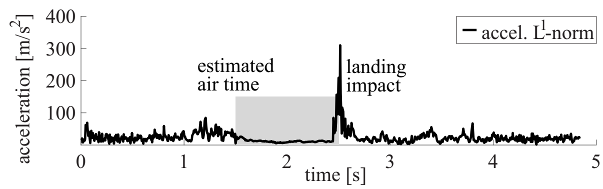
3) Parameter setting: The two-stage classification was based on six parameters: the time intervals $\Delta t_{rel,cat}$ and Δt_{pause}, the two thresholds in the magnetometer and accelerometer variance $\eta_{var,mag}$ and $\eta_{var,acc}$ and the weight factors of the classifiers $f_{w,grind}$ and $f_{w,air}$. The time intervals were defined by manual analysis of tricks of both acquisitions to be $\Delta t_{rel,cat} = 1$ s and $\Delta t_{pause} = 2$ s. The thresholds and weight factors were determined with a grid search during the classification process. The criteria for the parameter selection was based on the maximization of the F-measure [11].

E. Classification of trick class

An event that was determined as trick event of one category was further classified for the actual trick class. The



(a) Grind detection due to high magnetometer signal variance during the grind time on the metallic rail



(b) Air detection due to low accelerometer signal variance during the air time of the jump

Fig. 3. Example of magnetometer signal of grind trick (a) and accelerometer signal of air trick (b).

classification was performed separately for grinds and airs. However, both performances followed a similar procedure. The definition of the relevant signal period was based on the previously determined time of the landing impact. It was set to an interval from $\Delta t_{rel,trick}$ before the landing time to the landing time. In contrast to the category classification, the relevant signal period had to be set higher for the trick classification. In snowboarding, many tricks are already prepared before the actual grind time or air time. In order to include this preparation phase, the time interval was defined to be $\Delta t_{rel,trick} = 2$ s.

For the defined signal interval, features were extracted. These features were only based on the integrated gyroscope signal and only contained snowboard-specific rotation features. These were the total rotation per axis, the rotation per axis in the first half of the trick and the rotation per axis of the second half of the trick. Hence, nine features were calculated.

For the classification, four classifiers were compared: Naive Bayes (NB), C4.5, k-Nearest Neighbors (kNN), and Support Vector Machine with a radial basis kernel (SVM). kNN was evaluated for k $\in \{1,3,5\}$. For the SVM, a grid search was executed for the determination of the best performing parameters γ and C $\in \{10^{-6}, 10^{-5}, ..., 10^{6}\}$.

F. Evaluation

The evaluation of the event detection and subsequent classification in trick categories was based on both data acquisitions. These contained 275 tricks in total, 119 grinds and 156 airs. The event detection was evaluated by a leave-one-subject-out cross-validation. The results were calculated for the recall and precision of the detection of an actual trick event, without any differentiation of the trick category or class.

For the two-stage category classification of grind and air events, one Naive Bayes classifier was trained per classification step. The classifiers' performances were evaluated with a LOSO-CV. In addition, a grid search was performed for

the defined parameters of each stage. The classification with the best performing parameter configuration was evaluated by means of recall and precision.
Subsequently, the actual trick classification for both categories was evaluated based on the result of the previously performed event detection and category classification. Therefore, the trick classification could contain incorrectly detected and incorrectly categorized events. Both will be referred to as *no-trick* in the evaluation. The feature selection and classification with the four classifiers were performed with the Embedded Classification Software Toolbox (ECST) [12]. Note that the trick classification was only performed for data of acquisition B as stated in *Study design*. All performed tricks were manually labeled by analyzing the video recordings. All tricks were assigned to a trick category and for acquisition B in addition to a trick class.

III. Results

The evaluation of the event detection showed a positive detection of 274 out of 275 trick events and a false positive detection of 471 events. This results in a recall of 0.996 and a precision of 0.368.
The classification of trick categories was evaluated for the 745 previously detected events. The one trick event that was missed by the event detection will not be considered anymore. In the first stage, 115 out of 119 grinds were classified correctly. Furthermore, the grind classification contained 15 false positives. By the dead time considerations, 113 no trick events (but no actual trick event) were eliminated before performing the air classification. Here, 151 true positive and 15 false positive air events were detected while a total of four airs were missed in the category classification. Based on these numbers, the results were determined to a recall of 0.966 and a precision of 0.885 for grinds and a recall of 0.974 with a precision of 0.910 for airs.
The best performing classifier for the grind classification was SVM with an accuracy of 90.3 %. The air classification showed the best result for kNN with an accuracy of 93.3 %. An overview of the accuracies of all classifiers is provided in Tab. II. The confusion matrices with the best performing classifiers of both classifications are provided in Tab. III and Tab. IV.

TABLE II
RESULTS: ACCURACIES OF ALL COMPARED CLASSIFIERS FOR THE TRICK CLASSIFICATION OF BOTH CATEGORIES GRINDS AND AIRS.

accuracy [%]	NB	C4.5	kNN	SVM
Grind classification	87.6	86.7	89.4	**90.3**
Air classification	90.4	91.3	**93.3**	89.4

TABLE III
RESULTS: CONFUSION MATRIX OF THE GRIND CLASSIFICATION WITH SVM. THE MATRIX CONTAINS ALL CATEGORIZED GRIND TRICKS ACCORDING TO TAB. I AND THE *no-trick* CLASS OF INCORRECTLY DETECTED AND INCORRECTLY CATEGORIZED EVENTS.

	true			
pred.	50-50	BS-Boardsl.	FS-Boardsl.	no-trick
50-50	38	3	0	1
BS-Boardsl.	0	33	0	1
FS-Boardsl.	0	0	28	5
no-trick	0	0	1	3

TABLE IV
RESULTS: CONFUSION MATRIX OF THE AIR CLASSIFICATION WITH KNN. THE MATRIX CONTAINS ALL CATEGORIZED AIR TRICKS ACCORDING TO TAB. I AND THE *no-trick* CLASS OF INCORRECTLY DETECTED AND INCORRECTLY CATEGORIZED EVENTS.

	true			
pred.	Method	BS-180	FS-180	no-trick
Method	35	0	0	2
BS-180	1	32	0	2
FS-180	0	0	24	2
no-trick	0	0	0	6

IV. Discussion

The results of the event detection were evaluated to a recall of 0.996 and a precision of 0.368. The high recall results from the specifically set threshold ϵ that was defined by means of containing all trick events in the training phase. Although the majority of the incorrectly detected events could be excluded in the subsequent processing steps, a more sophisticated event detection approach could further improve the algorithm. The two-stage category classification performed effectively by incorporating one classifier per trick category and considering the dead time interval. However, a combination of the category classification with the event detection, which would already contain the dead time considerations between tricks, could widely simplify the processing chain. The trick classification showed results of more than 90 % for both trick categories. Considering the confusion matrices of both classifications, it can be seen that incorrectly classified events were mainly a result of the *no-trick* class. Without incorrectly detected and categorized events, the accuracy would be considerably higher.
One possibility to overcome this limitation would be a more thorough analysis of the time interval of the trick performance. In the proposed approach, the relevant time intervals for the category and the trick classification were manually defined to 1 s and 2 s before the landing impact. As a result, it could be possible that a too short or too long signal period was processed and thereby the events of one and the same category or class are less comparable. An improvement would be a determination of the actual grind or air time of the trick as suggested in [13], [14]. The determined duration of the trick would then be used for a classification based on an adaptive trick time interval.
With our approach, we established the first step towards an automated category and trick classification in freestyle snowboarding. However, we only analyzed two out of several categories. Further important categories are for example *Big Air* and *Half-pipe*. Although *Big Air* has

not been considered in this work, the basic idea of our air detection can be supposed to work as well if the aforementioned air time dependent trick time interval is incorporated. For half-pipe, the previous work of Harding et al. [2], [3] should be further analyzed and extended.

For the application in training sessions or competitions, two aspects have to be considered: the data transmission and the real-time capability of the system. While the current system was based on data which were stored on the sensor device, an application of the system would require a constant data transmission over the distance of the whole slope. Furthermore, the data transmission as well as the data processing have to be performed in real-time. Therefore, our algorithm was designed by means of low computational effort and short window lengths (e.g. in the event detection). The most crucial part of a real-time implementation is the incorporation of the proposed SVM classifier for the grind classification. However, in a final approach also alternative classifiers can be implemented and would still achieve comparable accuracies (see Tab. II).

V. Conclusion and Future Work

In this paper, we presented a full pipeline for the classification of multiple snowboard freestyle tricks. The evaluation showed a stable detection and classification rates of more than 90 % for all categories and tricks. The algorithm was designed to be extendable to more trick categories and trick classes. Hence, the proposed method provides the first step towards an application of trick classification in freestyle snowboard training and competitions. Furthermore, relevant information for the athletes, coaches and judges can be extracted from the processing pipeline. These include the already computed rotation angles and the proposed calculation of the trick duration.

The next steps towards an application in training sessions and competitions contain an extension of the trick categories in order to cover the full spectrum of freestyle tricks. In addition, further relevant parameters have to be defined for training support and attractive ways of motion visualization in competitions have to be developed. Furthermore, the algorithm has to be implemented in a real-time capable system. This includes the design of hardware components and the adaption of the algorithm for data transmission and processing in real-time.

Acknowledgment

The authors would like to thank Snowboard Germany, especially Nadine Härtinger, Andreas Jügelt and Friedl May as well as all participants of the study for their support during the data collection.

References

[1] J. W. Harding, K. Toohey, D. T. Martin, A. G. Hahn, and D. A. James, "Technology and half-pipe snowboard competition - insight from elite-level judges (P240)," in *The Engineering of Sport 7*, 2008, vol. 2, pp. 467–476.

[2] J. W. Harding, K. Toohey, D. T. Martin, C. Mackintosh, A. M. Lindh, and D. A. James, "Automated inertial feedback for half-pipe snowboard competition and the community perception," *The Impact of Technology on Sport II*, vol. 20, pp. 845–850, 2007.

[3] 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)," in *The Engineering of Sport 7*, 2008, vol. 2, pp. 447–456.

[4] T. Holleczek, J. Schoch, B. Arnrich, and G. Tröster, "Recognizing turns and other snowboarding activities with a gyroscope," in *14th Int. Symp. on Wearable Computers*, 2010, pp. 1–8.

[5] J. Anlauff et al., "A method for outdoor skateboarding video games," in *7th Int. Conf. on Advances in Computer Entertainment Technology*, 2010, pp. 40–44.

[6] B. H. Groh, T. Kautz, D. Schuldhaus, and B. M. Eskofier, "IMU-based trick classification in skateboarding," in *KDD Workshop on Large-Scale Sports Analytics*, 2015.

[7] 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*, 2014.

[8] F. Ferraris, U. Grimaldi, and M. Parvis, "Procedure for effortless in-field calibration of three-axial rate gyro and accelerometers," *Sensors and Materials*, vol. 7, no. 5, pp. 311–330, 1995.

[9] T. Holleczek, A. Rüegg, H. Harms, and G. Tröster, "Textile pressure sensors for sports applications," in *9th Annu. IEEE Conf. on Sensors*, 2010, pp. 732–737.

[10] P. Domingos, "Metacost: A general method for making classifiers cost-sensitive," in *5th ACM SIGKDD Int. Conf. on Knowledge discovery and data mining*, 1999, pp. 155–164.

[11] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *J. of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.

[12] 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*, 2012, pp. 2266–2269.

[13] 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*, vol. 6, no. 2, pp. 63–77, 2013.

[14] C. L. Roberts-Thomson, A. M. Lokshin, and V. A. Kuzkin, "Jump detection using fuzzy logic," in *IEEE Symp. on Computational Intelligence for Engineering Solutions*, 2014, pp. 125–131.