Recognition of Piano Pedalling Techniques Using Gesture Data

Beici Liang, György Fazekas, Mark Sandler
Centre for Digital Music
Queen Mary University of London
London, UK
{beici.liang,g.fazekas,mark.sandler}@qmul.ac.uk

ABSTRACT

This paper presents a study of piano pedalling technique recognition on the sustain pedal utilising gesture data that is collected using a novel measurement system. The recognition is comprised of two separate tasks: onset/offset detection and classification. The onset and offset time of each pedalling technique was computed through signal processing algorithms. Based on features extracted from every segment when the pedal is pressed, the task of classifying the segments by pedalling technique was undertaken using machine learning methods. We exploited and compared a Support Vector Machine (SVM) and a hidden Markov model (HMM) for classification. Recognition results can be represented by customised pedalling notations and visualised in a score following system.

CCS CONCEPTS

• **Information systems** → **Information retrieval**; • **Computing methodologies** → *Supervised learning by classification*;

KEYWORDS

Piano pedalling techniques, musical gesture recognition, machine learning in musical performance

ACM Reference Format:
Beici Liang, György Fazekas, Mark Sandler. 2017. Recognition of Piano Pedalling Techniques Using Gesture Data. In *Proceedings of AM '17, London, United Kingdom, August 23–26, 2017,* 5 pages.
https://doi.org/10.1145/3123514.3123535

AM '17, August 23–26, 2017, London, United Kingdom

ACM ISBN 978-1-4503-5373-1/17/08…$15.00
https://doi.org/10.1145/3123514.3123535

1 INTRODUCTION

Music performance is not only the realisation of the categorical pitch and duration information in the score, but it also involves an interpretation by the player, leading to expressive performance. An accurate transcription method should be able to detect the performance characteristics from both the intention of the composer written in the score and the interpretation of players, for instance, in the particular case of the piano, the use of pedals. Composers like Chopin and Liszt indicated the use of pedals actively in their work [11], while Debussy rarely notated pedalling despite the importance of pedals for the correct interpretation of his music [10]. Experts agree that pedalling in the same piano passage can be executed in many different ways even if pedal markings are given [3]. This is adjusted by the performer's sense of tempo, dynamics, textural balance, and the settings or milieu in which the performance takes place [10].

Modern pianos usually have either two or three pedals, among which the most frequently used is the sustain pedal. Using the sustain pedal is one of the main musical gestures to create different artistic expression in piano performance. It lifts all dampers and sets all strings into vibration due to sympathetic coupling. There are three main pedalling techniques considering the timing of the pedal with respect to the notes. Rhythmic pedalling is employed when the pedal is pressed exactly or very nearly at the same time as the notes. This technique supports metrical accentuation, which is an important aspect of Classic-era performance. Pressing the pedal immediately after the note attack is called syncopated or legato pedalling. This enables the performer to produce seamless legato. Anticipatory pedalling, first described in the 20th century, can only be applied after a silence and before the notes are played. This technique is used to produce greater resonance at the commencement of the sound. Besides variation in pedal timing, professional pianists apply other pedalling techniques that change as a function of the depth of the sustain pedal. These techniques include part-pedalling and fluttering pedal. Especially part-pedalling is used very often to colour the resonance subtly. Three or four levels of the sustain pedal are commonly defined in order to interpret the usage of part-pedalling.

No compositional markings exist however to indicate the variety of pedalling techniques mentioned above [12]. Although recognising and notating pedal use can benefit many applications, including automatic pedalling transcription or educational purposes, detecting pedalling parameters from the audio signal alone is a rather challenging task [4]. The purpose of this study is to devise a method for reliable recognition of pedalling techniques using gesture data collected with sensors from recorded piano performance. This can complement audio-based recognition or provide ground truth for developing algorithms that rely on audio only. The proposed method indicates the onset and offset times of the sustain pedal and the extent of the use of part-pedalling.

Machine learning can facilitate the analysis of expressive gestural features in music performance. This is primarily because of their ability to adopt to individual differences between players or interpretations, owing to their probabilistic representations of underlying signal data, or the ability to learn and exploit dependencies between the techniques employed [2]. Both Rasamimanana et al. [9] and Young [14] used a classifier based on the k-Nearest Neighbours (k-NN) algorithm to recognise violin bowing techniques. For piano playing gestures, Van Zandt-Escobar et al. [15] developed *PiaF* to extract variations in pianists' performance based on given gesture references, and use these estimated variations to manipulate audio effects and synthesis processes. However, the inclusion of pedalling techniques was not considered as part of gesture sensing in this or other related studies. Commercial computerised reproducing pianos such as Bösendorfer CEUS and Yamaha Disklavier have the ability to record continuous pedal position, but no intuitive representation of pedal usage is returned to users.

The approach taken in this paper follows the aforementioned ideas. Onset and offset time of pedalling gesture on the sustain pedal are firstly detected using signal processing methods. Two machine learning based methods (SVM and HMM) are proposed for classifying the segment between every onset and offset by four pedalling techniques (quarter, half, three-quarters and full pedal). Our pedalling technique recognition is demonstrated in an audio-based score alignment system, extended with markings of pedal use visualised in the context of the score.

The rest of this paper is organised as follows. We first describe the database that was built for our recognition task in Section 2. We then present the recognition process of pedalling techniques in Section 3. Recognition results of HMM and SVM classifiers are shown in Section 4. We conclude our work discussing potential applications in Section 5.

2 DATABASE CONSTRUCTION

Pedalling gesture of the sustain pedal can be captured using a new measurement system taken from our previous work [5].

Table 1: Number of pedalling instances in the music excerpts from our database.

Music Excerpts	1/4	1/2	3/4	full pedal
Op.10 No.3	14	13	7	5
Op.23 No.1	7	17	8	29
Op.28 No.4	17	24	5	24
Op.28 No.6	9	27	5	17
Op.28 No.7	2	10	3	1
Op.28 No.15	7	34	4	22
Op.28 No.20	9	12	11	17
Op.66	6	21	10	11
Op.69 No.2	2	15	10	24
B.49	3	51	8	17
Sums	76	224	71	167

In this system, near-filed optical reflectance sensing was used to measure the continuous pedal position with the help of Omron EESY1200 sensors. These were mounted in the pedal bearing block in order to avoid interference with pianists. The output voltage of the sensors was calibrated through a custom-built Printed Circuit Board (PCB) and recorded at 22.05kHz sampling rate using the analogue input of Bela[1], which is an open-source embedded platform for real-time, ultra-low-latency audio and sensor processing on the BeagleBone Black [6]. The piano sound can be synchronously recorded at 44.1kHz with the audio input to Bela as well.

The measurement system was deployed on the sustain pedal of a Yamaha baby grand piano. A pianist was asked to perform ten excerpts of Chopin's piano music using sheet music. Instructions describing the extent to which the sustain pedal should be pressed in each musical phrase were notated in the scores by the experimenter. The audio and the gesture data were then recorded to files. The gesture data was labelled according to the notated score in order to provide a basic ground truth dataset. Table 1 presents the number of instances of each pedalling technique in the music excerpts we recorded.

3 PEDALLING RECOGNITION

Our task is to recognise when and which pedalling technique were employed using the gesture data. This task was completed by using pedal onset and offset detection complemented by SVM- or HMM-based classification as described in the following sections. As we mentioned in Section 1, pianists vary their usage of pedalling techniques when the music piece or performance venue is changed. This requires calibration or automatic adaptation to how a techniques is used in a particular venue or a particular musician. Manually

[1] http://bela.io/

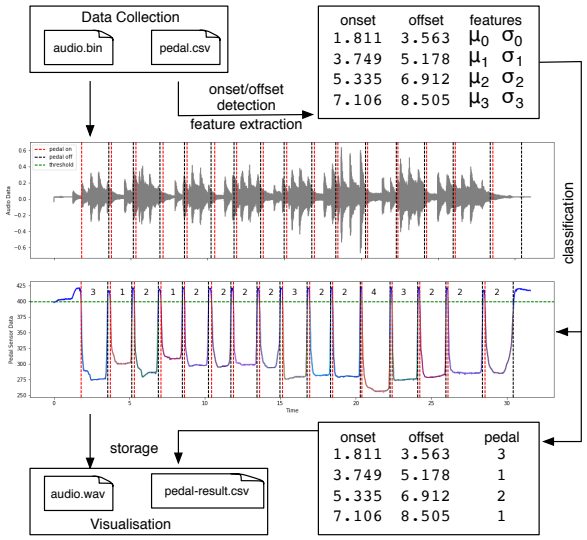

Figure 1: Recognition process.

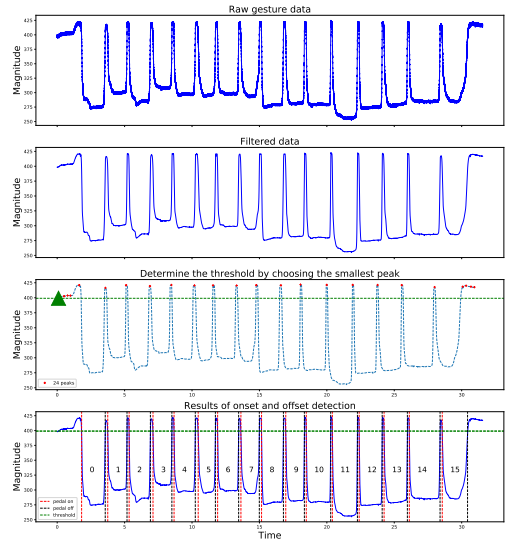

Figure 2: Onset and offset detection.

setting the thresholds to classify the level of part-pedalling is therefore inefficient. We decided to use the supervised learning to train SVM or HMM classifiers in a data-driven manner. To this end, we employed the scikit-learn [7] and hmmlearn[2] libraries to construct our SVM and HMM separately.

Figure 1 illustrates the overall process of recognition. Onset and offset times were determined using signal processing methods. Features were extracted from every segment which were defined by the gesture data between the detected onset and offset times. We observed that the shape of the histogram of the data in each segment fitted the normal distribution. Gaussian parameters were therefore extracted as features of the sensor signal for each instance of pedal use: μ is mean of the distribution and σ is standard deviation. Despite pedal position is measured and its use may be interpreted in a continuous space, classification of pedalling as discrete events and discrete types may benefit applications such as transcription and visualisation. We exploited an SVM and a HMM separately to classify the detected pedalling segments. A subset of our dataset was then used to train the classifier in order to output the label of the classified pedalling technique based on the features received from the remaining data. Label number 1 to 4 correspond to the quarter, half, three-quarters and full pedal technique. The recognition results maintained synchronised time with the audio file. They were then used as the inputs of our visualisation system.

Onset and Offset Detection

Figure 2 presents the process of onset and offset detection. The value of raw gesture data represents the position changes of the sustain pedal. The smaller the value, the deeper the pedal was pressed. It was first smoothed using a Savitzky-Golay filter, which is a particular type of low-pass filter well-adapted for smoothing noisy time series data [8]. This could avoid spurious peaks in the signal and the detection of false pedalling onsets or offsets. Based on the filtered data, pedalling onset and offset times were detected by comparing the data with a threshold (green dashed line). The threshold was selected by choosing the minimum value (green triangle) from a peak detection algorithm, i.e. the smallest peak. The moment when the value of data is smaller than the threshold is considered as the onset time (red dashed line), while larger is the offset time (black dashed line). In this way, each segment was defined by data between the onset and offset time. For example, there are 16 segments detected in Figure 2. Gaussian parameters of every segment were then extracted as the features for further classification.

Classification

The SVM algorithm finds the maximum margin hyperplane that separates two classes of data. If the data in the feature space are not linearly separable, they can be projected into a higher dimensional space and converted into a separable problem. For our SVM-based classification, we employed a SVM with linear kernel as a discriminative classifier to categorise one dimensional data into pedalling techniques. It essentially learns an optimal threshold for classification in a data-driven manner, avoiding the use of heuristic threshold.

The second method we employed and evaluated is HMM-based classification of time series data. A HMM is a statistical model that can be used to describe the evolution of observable events that depend on hidden variables which are not

[2]http://hmmlearn.readthedocs.org/

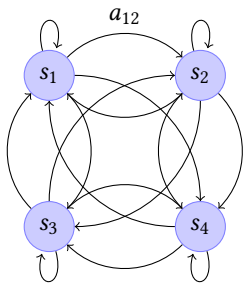

Figure 3: An HMM with 4 states representing four pedalling techniques. a_{ij} is the probability to transition from state s_i to state s_j.

directly observable. In our framework, the observations are the features from gesture data and the hidden states are the four pedalling techniques to be classified for our case. In our dataset which consists of Chopin's music, various pedalling techniques were used constantly. We assumed that learning the transition probability of the hidden states could reveal musicological meanings in terms of the extensive use of pedal for an expressive performance. We then employed our HMM model employed for classification. This was done by finding the optimal state sequence associated with the given observation sequence. The structure of our fully connected HMM with four states is shown in Figure 3, where states may exhibit self transition or transition into any of the three other states. We trained the probabilistic parameters for our HMM with Gaussian emissions. The hidden state sequence that was most probable to have produced a given observation sequence is discovered using Viterbi decoding.

Visualisation

In order to demonstrate a practical application of the above system, a piano pedalling visualisation system was developed that can present the captured data and the classification results to the pianist in the context of the score. a The visualisation system employed a score following implementation [13] in Matlab which aligns the musical score the audio recording of the piece. We extended this implementation to align the recognition results of pedalling techniques of the same piece given the detected onset and offset times and the pedalling technique. A screen shot of this system is shown in Figure 4. A music score has to be selected by the user first. After importing the audio recording and the corresponding pedalling recognition results of the same piece, they are displayed using the following markup: blue circles indicate the sounding notes in the score according to the audio, a star indicates pedal onset while the green square marks a pedal offset. The more dark red and lower a star is, the deeper the sustain pedal was pressed.

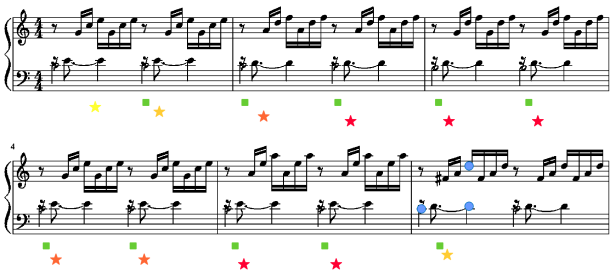

Figure 4: Screen shot of the visualisation system.

4 RESULTS

This present work focusses on the evaluation of pedalling technique classification. In terms of the quantity of our ground truth dataset, the performance of SVM and HMM were compared by conducting leave-one-group-out cross-validation. In this scheme, samples were grouped in terms of music excerpts. Table 2 shows the F-measure score of HMM and SVM classifiers respectively. Both classifiers were validated in each music excerpt where the data need to be classified, while the rest of the excerpts constitute the training set. A mean F-measure score of 0.801 and 0.925 was obtained for the HMM and SVM respectively.

Table 2: F-measure score of SVM and HMM.

Music Excerpts	HMM F-score	SVM F-score
Op.10 No.3	0.744	0.873
Op.23 No.1	0.902	0.967
Op.28 No.4	0.914	0.950
Op.28 No.6	0.759	0.906
Op.28 No.7	0.688	1.000
Op.28 No.15	0.627	0.919
Op.28 No.20	0.816	0.848
Op.66	0.938	0.938
Op.69 No.2	0.804	0.883
B.49	0.823	0.969
Mean	0.801	0.925

We suspect that the lower score of HMM is resulting from the fact that it was trained in a non-discriminative manner. The HMM parameters were estimated by applying the maximum likelihood approach, using the samples from the training set and disregarding the rival classes. Furthermore, a causality of one pedal position being followed by a certain other position may be unnecessary or adds very little value when the individual pedal events are separated from each other by long offset phases. For this reason, the learning criterion was not related to factors that may yield an

improvement of the recognition accuracy directly. The reported results can possibly be improved using the Hidden Markov SVM (HM-SVM) proposed in [1] as a discriminative learning technique for labelling sequences based on the combination of the two learning algorithms.

5 CONCLUSION

We presented the recognition results of piano pedalling techniques on the sustain pedal using gesture data collected using a dedicated sensor system. The temporal location of pedalling events were recognised using onset and offset detection through signal processing methods. The employed pedalling technique was then recognised using supervised machine learning based classification. Two classifiers using SVM and HMM were trained to separate the data into quarter, half, three-quarters or full pedalling technique. In our evaluation SVM outperformed the HMM based method.

The recognition results and the corresponding piano sound may be used in a score following application that visualises the performance parameters related to pedalling. This exemplifies a practical application of the system in music education or performance analysis. The techniques introduced in this study require sensors to be mounted on the piano which is not always practical, and may not help to analyse existing or historical recordings. Our future work includes the development of audio-based pedalling recognition techniques. This study can contribute to providing the ground truth dataset for detecting pedalling techniques from the audio alone. Evaluation of the visualisation system has not yet been conducted with users, which also constitutes future work.

ACKNOWLEDGMENTS

This work is supported by Centre for Doctoral Training in Media and Arts Technology (EPSRC and AHRC Grant EP/L01632X/1), the EPSRC Grant EP/L019981/1 "*Fusing Audio and Semantic Technologies for Intelligent Music Production and Consumption (FAST-IMPACt)*" and the European Commission H2020 research and innovation grant AudioCommons (688382). Beici Liang is funded by the China Scholarship Council (CSC).

REFERENCES

[1] Yasemin Altun, Ioannis Tsochantaridis, and Thomas Hofmann. 2003. Hidden markov support vector machines. In *Proceedings of the Twentieth International Conference on Machine Learning (ICML)*, Vol. 3. 3–10.

[2] Baptiste Caramiaux and Atau Tanaka. 2013. Machine Learning of Musical Gestures. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. 513–518.

[3] Elaine Chew and Alexandre RJ François. 2008. MuSA. RT and the pedal: the role of the sustain pedal in clarifying tonal structure. In *Proceedings of the 10h International Conference on Music Perception and Cognition, Sapporo, Japan.*

[4] Werner Goebl, Simon Dixon, Giovanni De Poli, Anders Friberg, Roberto Bresin, and Gerhard Widmer. 2008. Sense in expressive music performance: Data acquisition, computational studies, and models. *Sound to sense-sense to sound: A state of the art in sound and music computing* (2008), 195–242.

[5] Beici Liang, György Fazekas, Andrew McPherson, and Mark Sandler. 2017. Piano Pedaller: A Measurement System for Classification and Visualisation of Piano Pedalling Techniques. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME).*

[6] Andrew McPherson and Victor Zappi. 2015. An environment for submillisecond-latency audio and sensor processing on BeagleBone Black. In *Proceedings of 138th International Audio Engineering Society (AES) Convention.* Audio Engineering Society.

[7] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.

[8] William H Press, Brian P Flannery, Saul A Teukolsky, William T Vetterling, and Peter B Kramer. 1987. *Numerical recipes: the art of scientific computing.* AIP.

[9] Nicolas H. Rasamimanana, Emmanuel Fléty, and Frédéric Bevilacqua. 2005. Gesture analysis of violin bow strokes. In *International Gesture Workshop.* Springer, 145–155.

[10] Sandra P Rosenblum. 1993. Pedaling the piano: A brief survey from the eighteenth century to the present. *Performance Practice Review* 6, 2 (1993), 8.

[11] David Rowland. 2004. *A history of pianoforte pedalling.* Cambridge University Press.

[12] Deborah Rambo Sinn. 2013. *Playing Beyond the Notes: A Pianist's Guide to Musical Interpretation.* Oxford University Press.

[13] Siying Wang, Sebastian Ewert, and Simon Dixon. 2015. Compensating for asynchronies between musical voices in score-performance alignment. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 589–593.

[14] Diana Young. 2008. Classification of Common Violin Bowing Techniques Using Gesture Data from a Playable Measurement System.. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME).* 44–48.

[15] Van Zandt-Escobar, Baptiste Caramiaux, and Atau Tanaka. 2014. PiaF: A Tool for Augmented Piano Performance Using Gesture Variation Following. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME).*