

Response to review for submission ROBOT-D-15-00326

Dear Reviewers,

We would like to thank the reviewers for their time and valuable feedback. We appreciate all of the comments and suggestions, and believe addressing (and have addressed them in the revised version of the manuscript) them will make for a stronger paper. We hope that our text below will help to clarify the majority of the reviewers' concerns. We will address each one of the reviewers' concerns in a separate section.

Reviewer 1

1. *Page 2: The author stated that the proposed work in this paper aims to develop a generative model that captures the physics of a drum stroke (primarily focusing on multiple bounce strokes) as performed by a human drummer and implement the model in a robotic drummer through a proper controller design. However, in the section 2, the authors have omitted relevant related work (see below some examples) that is related to both physical model as well as robotic implementation. Therefore, the authors should clearly indicate the main technical contribution respect to the relevant related work. Furthermore, the author stated that none of the robot percussionists have been designed to produce double strokes or other kinds of expressive multiple striking techniques. However, the author should consider the work presented by Bicchi (2014), where double strokes with different stiffness settings have been realized by a robot system (actuator + physical model and control strategy).*

We appreciate you pointing us towards some work which we had not seen before. We have, as per your suggestions, enhanced the related works section with most of your recommendations. Reviewer 2 however had felt that “*related work is as comprehensive as the length of paper allows*”. Therefore, we have chosen a middle ground and including 4 out of 7 suggested related works.

We would like to clarify that the statement regarding double strokes was not a general one, but was specific to Weinberg's robots. However this line has been deleted in the final manuscript. We also have made the technical contribution clear in the paper.

2. The authors should provide further descriptions regarding the argumentations behind the choices while simplifying the model, e.g. why an inelastic collision has been considered? wouldn't be possible to consider the bouncing dynamics by considering an elastic collision)? Would be possible to express mathematically an approximation to the coefficient of restitution?

A. The collision between the drumstick and the drumhead is an inelastic collision because of the loss of kinetic energy due to damping, air resistance etc. The COR was just one of the parameters of the system and we chose to determine the approximate coefficient of restitution between the drumstick and the drumhead using experimental methods. Furthermore, the designer can be creative with the choice of values and experiment with it to create different kinds of stroke profiles. The focus of the work was not on finding a mathematical expression for COR (unlike in Bicchi's work), but to use the value of COR in the physical model when the simulations were performed.

3. The authors only indicate that a PID controller has been implemented but without giving details of the control system implemented, e.g. wouldn't be possible to include a block diagram of the control system? how the control gain parameters have been tuned?

A. A block diagram has been added which describes the flow of control and the framework implemented. Details regarding how the control gain parameters were tuned have been added to the text as well (the parameters were tuned manually by trial and error).

4. The authors should provide a short description of the mechanism for the percussionist robot so it is possible to understand the relation between the simplified model, the control system and the simulations.

A. We have added a short section in text describing the how the simulations and control system communicate with each other. The framework diagram, we hope, will also clarify this flow of information between different modules.

5. The authors should better describe the main objective for the simulation experiments presented. Furthermore, it is not clear how the simulation has been implemented. The simulation experiments should be discussed in detail and compare with other related works, e.g. Bicchi (2014), etc.

A. The main objective of the simulation experiments is to use the differential equation which describes the stick motion to generate trajectories (motor angle vs. time profiles) which can then be used as reference trajectories for PI control system. Clarifications have been provided in text regarding how the simulations were implemented. The work has been contrasted with the approaches taken in Bicchi in the related work. Our methodology drastically differs from a mass-spring damper model.

6. Wouldn't be possible to compare the bouncing frequency in a quantitative way (e.g. by means of FFT)?

The reviewer suggest to include a quantitative analysis (e.g. who the authors have tested the performance of the controller, what is the error achieved in terms of desired stroke frequency and the actual achieved one by the robotic drummer, etc.) before presenting the qualitative analysis.

A. We agree that it might be possible to measure and compare the bouncing frequency. But in this work we were not concerned about such objective metrics because for a multiple bounce stroke there is no one correct way to play the stroke and therefore establishing ground truth (and therefore the “desired” stroke frequency) for comparison is almost impossible. The work tries to understand the generative mechanisms from first principles and focuses on creating a perceptually indistinguishable aural effect. Therefore we chose to perform subjective tests for evaluations.

7.The authors should re-phrase this sentence for better understanding: “If the model was successful in emulating a human drum stroke completely, then the expected value for the number of incorrect answers would have been 5”.

A. The statistical analysis section has been revamped and the above mentioned statement has been rephrased and clarified.

8.The authors stated that “the statistical analysis of the results indicates that the generative physics model and its implementation in the RDP was successful in creating humanlike multi-stroke drumming expression”. However, it is rather difficult to understand the argumentations behind this statement from the physical point of view, e.g. robot side: desired bouncing frequency vs. detected bouncing frequency as well as human to robot comparison: detected bouncing frequency for the human performer vs. detected bouncing frequency for the robot performer, etc.

A. The statistical analysis has been revamped almost entirely as Reviewer 2 had pointed out a more fundamental flaw. Inferential confidence interval approach has been used in the current analysis. However the evaluation is still subjective as justified earlier (A. 6). And accordingly the conclusions have been revised as well.

9.The abstract and conclusion should be revised accordingly after revising the paper. The reviewer also suggest to re-consider a proper title that better describes the main content presented in this manuscript.

A. The abstract and conclusion have been revised. The title has been edited.

Reviewer 2

General notes:

I’m not sure the analysis of results allows them to draw the conclusion that the robot plays the drum strokes like a human (they use an inability to reject the null hypothesis in a t test as a sign of equivalence, whereas this just shows

indeterminacy). Accept if evaluation method can be clarified or corrected. Recommendations either to use a valid method for measuring statistical equivalence e.g. D. J. Schuirmann's (1987) use of 2 one-sided t tests or Tryon and Lewis's ICI approach (Tryon, W. W., and Lewis, C. (2008). An inferential confidence interval method of establishing statistical equivalence that corrects Tryon's (2001) reduction factor. *Psychological Methods*, 13(3), 272-277) or reanalyse the experiment so the hypothesis can be proved using a method of statistical difference where the rejection of the null hypothesis is a valid conclusion

A. We are very grateful for your comment regarding the flaw in the statistical analysis. We realize our mistake and learned quite a bit in process of revising the section. We have incorporated the statistical testing ideas mentioned in the papers that you had recommended. The entire statistical analysis section has been revised as per your suggestion and an inferential confidence interval approach was used to report a *maximum probable difference* between the two groups of data considered (data from the listening test and the simulated data). A test for statistical equivalence according to Tryon's method requires the specification of a Δ parameter which can be very arbitrary and meaningless in this context. Therefore we choose to just report the maximum probable difference which will still give an idea of how "similar" the two groups are.

2. Show a shorter range to get more detail in x axis of plots in section 4. I know they are that long to show the length of the thumb pressure envelope but we really can't see much variation in the strokes at the moment. Remove figure 6, its unnecessary

A. All the plots have been redone. The range has been reduced. The clarity of the figures has been improved. Figure 6 has been removed as well.

Reviewer 3

1. The authors propose a generalized physics based model for drum stroke generation for robotic drummers. The evaluation of the model is adequate and rigorous. Results show that the robotic drummer using this model is able to produce a number of strokes basically indistinguishable from human played strokes. Even though the average number of incorrect answers in the listening test is close to ideal, the standard deviation is quite high. The explanation provided in the paper is not fully convincing. Probably the authors in the future should better investigate this result.

A. The entire statistical analysis section has been revised as per request of Reviewer 2 and an inferential confidence interval approach was used to report a maximum probable difference between the two groups of data of considered (data from the listening test and the simulated data). The analysis in the revised manuscript is more sound than before.

We once again sincerely thank all the reviewers for their time and valuable suggestions and we look forward to the reply.

Sincerely,
Deepak and Gil