Using Wizard-of-Oz Techniques to Develop and Evaluate a Sonification System for Arm Motion in Breaststroke: A Small-Sample Pilot Study

Yueqiao Chen
Georgia Institute of Technology
Atlanta, GA
ychen3221@gatech.edu

ABSTRACT

A sonification for swimming can support the swimmers with real-time feedback on their strokes. However, most current swimming training relies on the coach's experience and video recordings, which can be difficult for swimmers to grasp fully. In this project, I built a swimming sonification system for advanced swimmers to train their arm motions according to their movement's accuracy and strength. The participants are three Georgia Tech students aged between 19-26 with at least three years of competitive swimming. In our study, we observed and evaluated the three participants interacting with the Wizard-of-Oz sonification to correct their arm motions in swimming. The research aims to try out the sonification system and test if the participants can identify and interpret correctly in the sonification. Experiments show that participants can understand each swimming phase and accuracy most of the time in sonification. The sonification can improve the accuracy of the arm motions. Different noises in sonification can influence the interpretation of sonification. The participants are willing to try sonification in future training but do not feel the sounds enjoyable.

ACM Computing Classification System

CCS → Human-centered computing → Human computer interaction (HCI) → Interaction devices → Sound-based input / output

Keywords

Swimming; sonification; breaststroke; stroke technique; voice training assistants; training

1. INTRODUCTION

Swimming is a popular sport that requires a high degree of technical proficiency to excel. One key component of swimming technique is the proper use of arm motions. However, many swimmers struggle with developing correct form, leading to inefficient swimming and reduced performance. Mid and advanced-level swimmers would like more precise gesture detection during their swimming. Existing training strategies often rely on a coach's subjective observations and verbal cues, but the coach could be mistaken or miss some errors for the swimmers. The swimmers received feedback after many laps couldn't recall which stroke they did wrong. That's why it is necessary to build real-time feedback on swimming for mid and advanced-level swimmers. Therefore, the goal of the project is to provide swimmers with the real-time sonification of the accuracy and force of their arm motions in the breaststroke to help train their stroke technique. In this investigation, I focus on the swimming domain: breaststroke and its arm motions.

Our primary research question is ***(RQ1) How accurately can swimmers identify and interpret the sonification of their stroke kinematics?*** I answer this question by collecting my wizard-of-oz data and the self-report data for each participant's trial. By comparing my wizard-of-oz data with self-reported accuracy and force, we could determine the difference between their interpretation and the data I predicted. If the average difference for a participant is over 0.1, it is a significant error. I also have survey questions asking how they think they could interpret the data in sonification. ***(RQ2) Can they differentiate the sound from the background noise and water sound?*** I answer this question via a controlled experiment. The participant was asked to perform four similar breaststroke arm motions for each trial: one without sonification, the second with sonification but no noise, the third with sonification and water sound, and the last one with sonification with a reverb effect. The Wizard-of-oz and self-report data are collected to compute a difference that indicates the deviation of interpretation and data I predicted. I also make survey questions to ask how much they think water sound and reverb sound effects interfere with sonification. ***(RQ3) What are the swimmers' attitudes toward the use of sonification during swimming practice? Do they feel the sonification enjoying or annoying?*** This part of the data is collected through surveys. The participants will give a score between 1 and 10 to show the extent of agreement with the statement. If the average of 3 participants on a question is above 6, the attitude toward the question is considered positive.

2. METHODS

2.1 Participants

Participant A is a female, aged 19, with three years of competitive swimming. Participant B is a female, aged 26, experiencing five years of competitive swimming. Participant C is a 23-year-old male with six years of swimming. All the participants are mid or advanced-level swimmers and familiar with breaststroke. The participants include a person aged under 20, between 20 and 25, and a person older than 25. The variety of ages and gender can provide a more diverse and less biased dataset.

2.2 Setting

The study is conducted at the CRC swimming pool. The swimming pool provided a much more realistic environment to conduct our experiments and listen to the sonification.

2.3 Systems

The sonification system can generate sounds according to the JSON data or use the volume slider to monitor the phases and strength of their arm movements.

2.3.1 JSON Data Simulation

2.3.1.1 JSON Data variables and ranges

The JSON data contains a list of Notification objects, each representing an arm motion. Each notification contains phase, force, timestamp, duration, oxygen, and accuracy data. In breaststroke, the arm motions are divided into five phases: glide, out-sweep, catch, in-sweep, and recovery. Table 1 shows each phase's usual range of force and duration.

Table 1. Range of Force and Duration for Each Phases of Arm motions in Breaststroke

Phase	Force (N)	Duration (ms)
Glide	≤10	1000 ~ 2000
Out-sweep	50 ~ 150	200 ~ 400
Catch	100 ~ 300	500 ~ 700
In-sweep	150 ~ 400	500 ~ 700
Recovery	20 ~ 50	200 ~ 400

The oxygen level range should be above 94 percent for trained swimmers, or it could cause danger. The accuracy of arm motions is usually above 0.6 to maintain swimming. The data objects in the data documents are usually taken in the given ranges of features.

2.3.1.2 Sonification Control Based on JSON Data

The sonification system can generate according to the timestamp, force, duration, oxygen, and accuracy as follows:

- The timestamp is the cumulative duration of the swimming and is used to schedule tasks.
- A sin wave is controlled by force and duration. The duration and force/350 as gain value are passed in the amplitude/gain envelope to manipulate the gain of the sin wave.
- If the oxygen level is under 94%, an alarm sound is raised, like "D, D."
- The accuracy controls another square wave in the sonification. The accuracy * 440 gives the frequency of the square wave.

2.3.2 Sonification Controls on Researchers (Wizard-of-Oz)

There is a volume slider from 0 to 1 for the researcher to control the wave volume. The researcher can use the example 2 JSON data as a template, which presents the sonification of average swimmers with assumed constant accuracy at 1. After the sampled data sonification was presented and researchers understood how to control the slider according to arm motion, a researcher, I, can modify the volume according to the phases and arm forces of the participants.

Initially, the researcher should move the volume slider to 0 so the swimmers hear nothing. After the example 2 sonification ends, the researcher can start their control. The control of volume is according to the force and phase of the participants present. For example, the glide doesn't need much force, so the sonification is low volume. The maximum glide value in the volume slider is 1.0. The researcher gives the 1.0 volume corresponding to 350N or higher force, usually in the In-sweep phase.

2.3.3 Participant Interfaces and Equipment

Participants were asked to perform arm motions while standing on dry land. The participants are equipped with earbuds to listen to the sonification while performing the stroke on dry land.

2.4 Procedure

2.4.1 Training

Participants will be informed about the purpose of the evaluation, what the study involves, and their rights as participants. Participants will receive an introduction to the swimming simulator and the sonification feedback. I will provide the three sonification examples to familiarize them with the system. They will also be told about the range of force and accuracy, and I will perform some example arm movements with force and accuracy being told.

2.4.2 Experiments

Each participant was asked to perform four individual arm movements: one without sonification, one with sonification but no noise, one with sonification with water sound, and one with sonification with a reverb effect. For each arm movement, they are asked to self-report their movement accuracy. The detailed tasks are provided in Table 2.

Table 2. Tasks

Task No.	Instructions
Task 1	Participants will make the arm motion on dry land without the sonification in earbuds for a stroke. I will record my predicted accuracy for the stroke. The participants should self-report the accuracy of the stroke.
Task 2	Participants will make the arm motion on dry land with the sonification in earbuds for a stroke. I will record my predicted accuracy for the stroke. The participants should self-report the accuracy of the stroke.
Task 3	Participants will make the arm motion on dry land with the sonification and water sound in earbuds for a stroke. I will record my predicted accuracy for the stroke. The participants should self-report the accuracy of the stroke.
Task 4	Participants will make the arm motion on dry land with the sonification with the reverb sound effect in earbuds for a stroke. I will record my predicted accuracy for the stroke. The participants should self-report the accuracy of the stroke.

2.4.3 Post-experiment Surveys

After the experiments, they are asked to complete a survey about the usability and satisfaction of the sonification system. The sample questions are presented in Table 3.

Table 3. Survey Questions

Question No.	Questions
Q1	*On a scale from 1 to 10, how do you agree to use sonification in training?*
Q2	*On a scale from 1 to 10, how do you score the effectiveness of sonification in training?*
Q3	*On a scale from 1 to 10, how do you score the effect of water sound on the quality of sonification?*
Q4	*On a scale from 1 to 10, how do you score the reverb effect influence on the quality of sonification?*
Q5	*On a scale from 1 to 10, how accurately could you interpret the phases of the arm motions in breaststroke?*
Q6	*On a scale from 1 to 10, how much do you agree that sonification is not annoying?*

2.5 Measures

I collect the accuracy of the interpretation according to self-report. The predicted accuracy was measured and collected by me according to the Wizard-of-Oz experiments. I will compute the difference between the two streams of data and then average for each condition to see the difference between interpretation and actual sonification. The surveys are all scored in the range of 1 to 10, so I will average the score from 3 participants to check their attitude on each statement. If the rounded average is higher than 6, it is an overall positive attitude.

2.6 Shortcomings

My prediction accuracy is not very accurate or precise. If we have sensors that can detect the postures precisely and calculate the accuracy, the accuracy will be more meaningful. It would be more indicative of asking the participants to swim with sonification, but I don't have waterproof Headphones provided. Also, three participants are a small sample. The data may not accurately represent the population of interest, leading to biased or unreliable results. It is possible to ask them to perform more trials of experiments. In this study, I only present one trial and its data.

3. RESULTS

Below is the raw data of the experiments on each stroke the participants did. Table 4 is the raw data**.** ***(RQ1)*** and ***(RQ2)*** will use the data to figure out the interpretation accuracy of the sonification.

Table 4. Raw Data

Sonification Type		**Participants**		
		A	**B**	**C**
Without Sonification	**Self Report**	0.6	0.8	0.9
	Predicted	0.7	0.7	0.9
With Sonification (No Noise)	**Self Report**	0.8	0.8	0.7
	Predicted	0.8	0.9	0.9
With Sonification and Water Sound	**Self Report**	1.0	0.9	0.9
	Predicted	0.9	0.8	0.8
With Sonification and the Reverb Effect	**Self Report**	0.9	0.7	0.6
	Predicted	0.8	0.9	0.8

Table 5 shows the average accuracy in different conditions of sonification for self-report accuracy and predicted accuracy. ***(RQ1)*** and ***(RQ2)*** can use Table 5 to find the accuracy improvement while using sonification or sonification with noises.

Table 5. Average Accuracy

Sonification Type		**Average Accuracy**
Without Sonification	**Self-Report**	0.77
	Predicted	0.77
With Sonification (No Noise)	**Self-Report**	0.77
	Predicted	0.87
With Sonification and Water Sound	**Self-Report**	0.93
	Predicted	0.83
With Sonification and the Reverb Effect	**Self-Report**	0.73
	Predicted	0.83

For each trial, we calculate the difference between the self-report accuracy and the accuracy I predict. The equation is:

$$difference = self\_report - predicted.$$

Then I average the difference of the 3 participants for each condition to see the differences between interpretation and predicted data in Wizard-of-Oz sonification. Table 6 shows the average difference in the four conditions.

Table 6. Average Differences

Sonification Type	Difference Between Self-report and Predicted for Each Participants			Average Difference
	A	B	C	
Without Sonification	-0.1	0.1	0	0
With Sonification (No Noise)	0	-0.1	-0.2	-0.1
With Sonification and Water Sound	0.1	0.1	0.1	0.1
With Sonification and Reverb Effect	0.1	-0.2	-0.2	-0.1

Table 7 shows the scores each participant gave in the surveys for each question. ***(RQ1)***, ***(RQ2)*** and ***(RQ3)*** can use the data. ***(RQ1)*** use it to find whether participants can interpret the phases correctly according Q5. ***(RQ2)*** will check Q3 and Q4 to see whether the participants think noise, like water sound, and the reverb sound effect could affect the quality of the sonification. ***(RQ3)*** will see the usability and satisfaction of the sonification and attitude according to Q1, Q2, and Q6.

Table 7. Survey Data

Question No.	Participants		
	A	B	C
Q1	10	9	9
Q2	7	8	8
Q3	7	5	4
Q4	6	3	8
Q5	9	10	8
Q6	7	6	6

Then, I average the score for each question. Table 8 presents the average score for each question.

Table 8. Average Score for Each Questions

Question No.	Average Score
Q1	$9.33 \approx 9$
Q2	$7.67 \approx 8$
Q3	$5.33 \approx 5$
Q4	$5.66 \approx 6$
Q5	9
Q6	$6.33 \approx 6$

4. DISCUSSION

4.1 Results' Findings

(RQ1) How accurately can swimmers identify and interpret the sonification of their stroke kinematics?

In Table 4, I compared the predicted accuracy without sonification and with sonification (no noise) conditions. It is shown that Participants A and B improved their predicted accuracy of arm motion, and Participant C remained at 90% accuracy. In Table 5, we can see the same self-report average accuracy with and without sonification but a higher average accuracy with sonification in predicted accuracy. An improved predicted accuracy indicates a potential improvement while using the sonification in arm motions. According to Table 6, the average difference between self-report and predicted accuracy without sonification is 0, while that with sonification is -0.1. It illustrates that participants with sonification intend to under-consider their performance, which means the self-report accuracy is average less than the predicted accuracy. According to data in survey Question 5 about the interpretation of swimming phases, the average scale participants gave is 9 in Table 8, and all participants gave a scale over 6, so all participants think they could interpret the swimming phases of arm motions in sonification.

Based on the data, the results show that participants can interpret swimming phases and have an initial understanding of the accuracy of their arm motions. Using sonification can improve the accuracy of their arm motions, but the sample is not large enough to ensure the improvement. The results also found that participants underestimate their accuracy while using the sonification. It is possible because the accuracy of their arm motions improved during sonification, but the participants didn't feel about that.

(RQ2) Can they differentiate the sound from the background noise and water sound?

According to the predicted accuracy in Table 4 and the average predicted accuracy in Table 5, the participants still improved their performance, but not as well as the sonification without noise. In Table 6, it is Participants hearing sonification with water sound intended to give a higher self-report accuracy, while hearing sonification with the reverb effect leads to underestimation of their accuracy of arm motions. It could be because the water sound gives them a much more familiar feeling of swimming, so they think they performed better. The underestimation regarding the reverb effect could result from the damping in the reverb effect and give them a blurrier hearing. Table 8's Q3 and Q4 provide an average scale under or equal to 6, which means that participants did feel about the influence of the water sound and reverb effect, but they may overcount the impact of the noise in self-report accuracy.

(RQ3) What are the swimmers' attitudes toward the use of sonification during swimming practice? Do they feel the sonification enjoying or annoying?

In Table 3, Q1 shows the attitude of participants toward the usability of the sonification, Q2 shows the effectiveness of the sonification they thought, and Q6 asks whether they are enjoying the sonification or not. Table 8 gave an average scale on Q1 and Q2 above 6, which means that participants are positive about the usability and effectiveness of the sonification in swimming training. The average scale of Q6 is about 6, which shows that current sonification may sometimes sound annoying to the participants but is somewhat helpful in training.

4.2 Strength, Weakness and Future Iterations

The study's strength is that it used a controlled experiment to consider the noise impact on sonification conditions. It will give a more realistic and practical result for considering variables influencing sonification's effectiveness. The weakness, however, is obvious. I did not have a large sample for a more representative and diverse result. The equipment is simple, and equipping waterproof headphones can bring more meaningful results. There is only one trial of the experiment, which could cause many biases.

Although there are some drawbacks, I think the data I collected is enough to give initial feedback on the sonification I made. The average accuracy and difference accuracy for the four conditions of sonification gives a good illustration and evaluation of the functionality of sonification, as discussed in RQ1 and RQ2. The survey data and its average scale provide a subjective attitude of the participants toward the sonification system, as mentioned in RQ3. Therefore, the data collected is sufficient to answer the research questions. For future iterations of research, I shall find more participants and experiment with more trials to have a less biased dataset. I would also buy waterproof headphones to give the participants actual swim with the sonification.

5. CONCLUSION

In conclusion, the sonification of swimming could help adjust the arm motions in training. The noise could affect the quality of sonification but not too much. The swimmers' perception of sonification could vary with different noises, but they could still reasonably interpret the swimming phases and accuracy. But the participants feel the sonification is somewhat annoying, so it is possible to use other sound waves or sound effects instead of sin waves to simulate the sonification. The next step could be simulating sonification using different sound waves and evaluating their enjoyment of different simulations.