C. ELEGANS MEETS DATA SONIFICATION: CAN WE HEAR ITS ELEGANT MOVEMENT?

Hiroko Terasawa[1], Yuta Takahashi[1], Keiko Hirota[1], Takayuki Hamano[2], Takeshi Yamada[1], Akiyoshi Fukamizu[1], Shoji Makino[1]

Life Science Center of TARA, University of Tsukuba, Ibaraki, Japan.[1]
JST, ERATO, Okanoya Emotional Information Project, Tokyo, Japan. [2]
terasawa@tara.tsukuba.ac.jp, y-takahashi@tara.tsukuba.ac.jp, hirota@tara.tsukuba.ac.jp, hamano@japan.com, takeshi@cs.tsukuba.ac.jp, akif@tara.tsukuba.ac.jp, maki@tara.tsukuba.ac.jp

ABSTRACT

We introduce our video-data sonification of *Caenorhabditis elegans (C. elegans)*, a small nematode worm that has been extensively used as a model organism in molecular biology. *C. elegans* exhibits various kinds of movements, which may be altered by genetic manipulations. In pursuit of potential applications of data sonification in molecular biology, we converted video data of this worm into sounds, aiming to distinguish the movements by hearing. The video data of *C. elegans* wild type and transgenic types were sonified using a simple motion-detection algorithm and granular synthesis. The movement of the worm in the video was transformed into the sound cluster of very-short sine-tone wavelets. In the evaluation test, the group of ten participants (from both molecular biology and audio engineering) were able to distinguish sonifications of the different worm types with an almost 100% correct response rate. In the post-experiment interview, the participants reported more detailed and accurate comprehension on the timing of the worm's motion in sonification than in video.

1. INTRODUCTION

1.1 Background and Goal

One of the most promising directions in data sonification is the sonification of time-series data, because auditory perception is very sensitive to changes in time [1, 2]. EEG data sonification and seismic data sonification offer successful examples that intuitively display transitions over time, inspiring sonification of other kinds of dynamic, time-series data [3, 4]. In disciplines such as biology, researchers observe organisms by visualization or quantitative measurements, and sonification has seldom been applied as a data-observation technique. However, biological research investigates temporal change of life and organisms, and we expect that sonification may provide a means to comprehend biological phenomena from a new angle.

Molecular biologists are currently experiencing major advances in their research methods. The development of high-resolution video recording and the usage of fluorescent protein tags have greatly expanded the possibilities for *in-situ* observation (i.e. non-destructive, real-time observation of living organisms), leading to high expectations for new discoveries by observing dynamic motions. Although such dynamic data are currently observed by "eyeballing" video data, we expect that sonifying temporal elements from video data could be equally advantageous to visual display, offering another modality of data observation in molecular biology.

Since sonification is still a new and unconventional approach for many people outside the auditory display community, we suspect suddenly switching to an abstract auditory display might seem initially implausible to biologists. In order to gain acceptance of sonification as a convincing research method, we need simple and straightforward sonification examples, where the causality between the original data and the resulting sounds can be easily grasped. Therefore, we employed video recording, which is often assumed to be the most concrete visual data, for our preliminary investigation in sonification. Our goal in this study is to examine whether we could "hear the movements we see" in a transparent manner, and thus to motivate further research directions.

1.2 Dynamic Movements of Model Organisms

Caenorhabditis elegans (C. elegans) is a small nematode, only about 1 mm long and formed from precisely 959 somatic cells. In the 1960s, Sydney Brenner began using the tiny worm to study the genetics of development, and it has since been used extensively as a model organism [5]. Brenner, John Sulston, and Robert Horvitz shared the 2002 Nobel prize in physiology or medicine for their discoveries in *C. elegans* concerning genetic regulation of development and programmed cell death.

C. elegans has been a popular model organism because of its favorable characteristics for biological experiments. Various genetic and biological techniques have been invented, enabling experiments on development, the nervous system, and aging/longevity.

1.3 Sonification of Image and Motion

"What we see" in video data is an image in motion, and the sonification of a moving image poses its own unique challenge, in contrast to the sonification of a static image or of motion. One common approach in the sonification of a static image is to unfold the image data into a raster sequence of pixel data and to interpret it as a waveform representation of sound [6, 7]. However, this method does not translate the visual sensation of up/down and left/right into audio in an intuitive way, and information about the object position is lost. Meanwhile, sonification of motion is often approached using acceleration sensor data or motion-capture data, at pre-selected measuring points [8, 9]. Substituting sensors and motion capture with video data has also been proposed, such as in human gait sonification, for which Boyd *et al.* extracted a phase configuration that describes the timing pattern of motions in the gait and sonified that data [10].

As that example indicates, video data represent raw image data, and so it is necessary to computationally extract meaningful and intuitive information about visual objects (such as size, speed, position, etc.) and eliminate what is unnecessary background image. Pelletier discussed the matching between visual object and sound object and proposed a framework to sonify the vector representation of corner displacement, which is a perceptually salient feature in vision [11]. This system could employ any kind of synthesis method, but the use of granular synthesis [12], which can represent the addition of large-number simple components, is suggested as one of the natural choices.

1.4 Framework of the Study

In this work, we sonified video data of *C. elegans* wild type and transgenic types. Worm movement in the video was transformed into sound cluster, using a simple motion-detection algorithm and granular synthesis. We examined the resulting sounds with an evaluation test, in which both biologists and audio engineers participated. The effect of sonification was measured with an identification task, in which the participants judge which video the presented sound was generated from. In the next sections, we describe the genetic manipulation, sonification method, and evaluation test, followed by discussion and ideas for further research.

2. GENETIC MANIPULATION OF *C. ELEGANS*

In this study, we investigated three kinds of *C. elegans*, the wild type, the red-fluorescent type, and the rolling type. The latter two types were transgenic strains, and they were generated using standard microinjection methods [13], in which DNA solution was injected into worm gonads.

Wild type: Bristol N2 wild type was provided by the Caenorhabditis Genetics Center [14].
Red-fluorescent type: We prepared transgenic worms expressing red fluorescent protein in their pharynxes (the worm's throat) by injecting DNA (promoter myo-2::DsRed plasmid) into wild-type worms [15].
Rolling type: To generate transgenic worms displaying the rolling movement phenotype, the plasmid containing the rol-6 (su1006) gene was injected into wild-type worms [16].

C. elegans was grown on E. coli lawn (as food), on agar plates. Video recording was done using a Leica MZFL III microscope and Leica DFC500 digital camera with a resolution of 1168x878 pixels.

Out of these recordings, we prepared four videos (A, B, C, and D) of 20-second duration each. Table 1 shows the list of worm types and video data, and Fig. 1 shows the screenshots of the videos.

Type of Worm	Video Data
Wild	A, C
Red-fluorescent (transgenic)	B
Rolling (transgenic)	D

Table 1. The List of *C. elegans* Types and Video Data

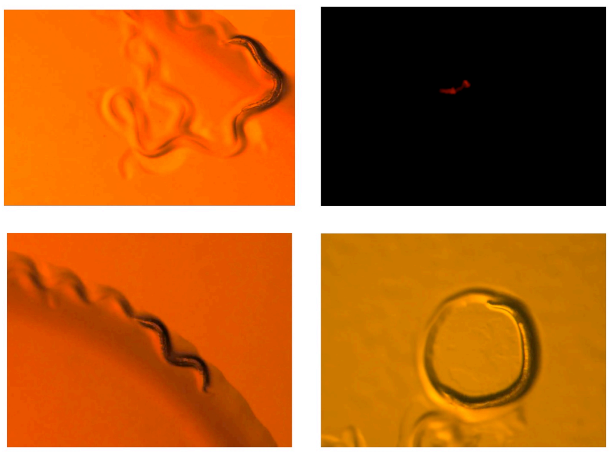

Figure 1. The snapshots of the video data (A: top-left, B: top-right, C: bottom-left, D: bottom-right).

3. SONIFICATION DESIGN AND SOUND SYNTHESIS

3.1 System Overview

The algorithm for sonification was implemented using Max/MSP/Jitter [17] as shown in Fig. 2. We apply a simple motion-detection algorithm on the video data to extract the moving worm from the background, then the image resolution is rescaled to 80 x 60 pixels. The downsized video shows a rough figure of the worm with a cluster of pixels. Granular synthesis translates the cluster of pixels into a sound cluster.

3.2 Motion Detection

The worm is filmed on an agar plate and exhibits some background objects such as traces of the earlier movements. In order to extract the moving worm, the system reads the video frame every 40 milliseconds. The absolute difference between the current read-out frame and the prior read-out frame is calculated, enhanced by raising the value to the fourth power. Then the extracted motion is smoothed out using envelope-following. The resolution is rescaled to 80 x 60 pixels to minimize the dataflow, and thus the computational power required for granular synthesis.

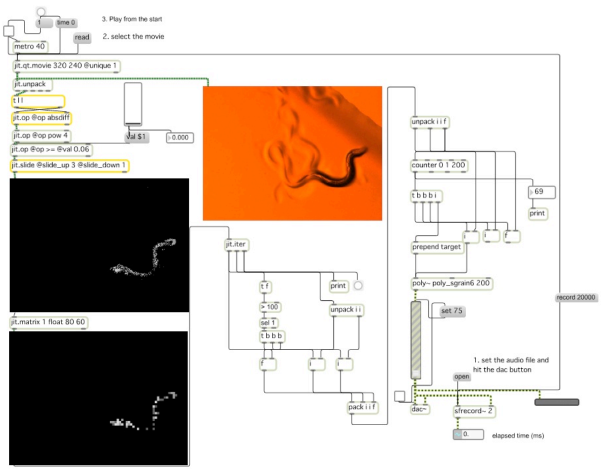

Figure 2. Max/MSP/Jitter implementation of the sonification algorithm.

3.3 Granular Synthesis

The extracted worm image is sonified using the granular synthesis technique, since the granular synthesis is suitable to capturing pixel representation of the image and reflects the complexity of the image into sound. For each pixel of the worm, a small particle of sine wave (wavelet or sound grain) was generated using the parameters of horizontal and vertical positions in the frame (x and y axes) and the intensity of the pixel value. The cluster of the pixels for the entire worm is heard as the sum of all the corresponding wavelets.

The mapping for the granular synthesis was decided upon by trying several configurations. We designed the vertical axis to correspond with the pitch, the horizontal axis to correspond with the attack-time and duration, and the pixel value to correspond with the intensity of the wavelet. All of these acoustical characteristics are designed to vary exponentially based on the perceptual scaling [18, 19].

4. EVALUATION TEST

4.1 Procedure

We conducted an evaluation test to judge the effect of the sonification using an identification task. We designed the task to resemble the expected use, in which the users are well informed about the sonification concept and the algorithm.

Four sounds (A, B, C, D) were created from the four video-data sources (A, B, C, D, respectively). Then four video clips with matching sounds (i.e., video A and sound A) were produced. These videos are available on our website:
http://www.tara.tsukuba.ac.jp/~terasawa/Worms/SMC2011.htm

The participants first received an explanation of the sonification algorithm by watching the Max/MSP/Jitter patch, and then they proceeded to the practice session, in which they watched the four video clips together with sound a few times to familiarize themselves with the sound. In the test session, the participants then listened to the sounds only and were asked to identify the video from which the sound was generated. During the test session, we provided a card showing the snapshots of the video with the names (A, B, C, D), so that the participants did not need to memorize the video names. The participants were allowed to listen to the stimuli several times if they wished.

Each stimulus was presented four times, resulting in 16 stimulus presentations. The stimuli were presented in a randomized order, but ensuring that any particular stimulus would not appear twice in succession. The chance level for making a correct response at each trial is 25 %.

4.2 Participants and Environment

Five molecular biologists (three graduate students and two faculty members) and five audio engineers (two graduate students and three faculty members) participated in the experiment.

The test was conducted in a normal laboratory office space (i.e., not particularly quiet) to resemble realistic user conditions. The stimuli were presented with built-in audio of a laptop computer (Apple MacBook Air) with a closed-type stereo headphone (Sony MDR-7506). The participants adjusted the volume to a comfortable level.

4.3 Questionnaire and Interview

After the test session, we asked four yes/no/maybe questions. The questions were:

(1) Is it easy to associate the sound and video?

(2) Can you hear the change of the worm's position?

(3) Can you hear the rhythm of the worm's movement?

(4) Do you think you will improve your hearing with more practice?

After the participant answered these questions, the experimenter had a free-form interview with the participant to gather useful comments and suggestions.

5. RESULTS

5.1 Evaluation Test

The mean correct response rate from the identification task is shown in Table 2, which was calculated by averaging the percentage of correct responses across the participants.

All the molecular biologists performed the identification task perfectly. Audio engineers performed slightly less well, but still above 95% correct. Overall, the participants performed the identification task very accurately.

Group	Mean Correct Response Rate
All	98.2%
Molecular Biologists	100%
Audio Engineers	96.4%

Table 2. The Percentage of Correct Responses in the Identification Task

5.2 Questionnaire

The percentage of "yes" answers for the questionnaire is shown in Table 3. The percentage was calculated by taking the sum of responses by counting "yes" as 1, "no" as 0, and "maybe" as 0.5, divided by the number of participants.

With question 1, molecular biologists and audio engineers showed different attitudes with their confidence in hearing. This difference may be because audio engineers are more used to working with sounds, or that there were more musicians among the audio engineers.

With questions 2 and 3, both molecular biologists and audio engineers provided the same type of responses. Most of them recognized the change in position, but they were not confident that they heard the "rhythm" of the movement. We asked question 3 expecting that the participants could identify some patterns of the movement. Perhaps the use of the word "rhythm" was not appropriate because it implies very periodic patterns, while the worm shows only pseudo-periodic patterns with its movement.

With question 4, all of the participants answered that they could improve the hearing by practice. The participants showed almost 100% accuracy in the identification task. Improving the hearing would lead to the easier comprehension of the sounds, if the accuracy is already accomplished.

Question #	Molecular Biologists	Audio Engineers
1	40%	90%
2	100%	90%
3	60%	60%
4	100%	100%

Table 3. The "Yes" Answer Percentage of the Questionnaire

6. DISCUSSION

During the interview, many of the participants stated that the main identification cues were the density of sound (i.e., the amount of wavelets) and the pattern of pitch change (i.e., vertical displacement).

The density of sound becomes very low with the red-fluorescent type. Because the number of visible cells in the worm's body is very few, only a small number of wavelets exist in the sound. With the other types, the density of sound is low when the worm is stopping or showing only tiny movements. The participants used the timing of sparse sounding and silence as a cue to identify the movement patterns.

The biologists also reported that they were able to focus on the worm movements more accurately with sound than with video. There are some moments when the worms briefly stop their motion. With sounds, such moments are easily detected. But with video, the participants tend not to notice such moments, and have the perception that the worms are moving smoothly without any interruption.

Both engineers and biologists reported they used the density and pitch cues concurrently. However, the parameters that corresponded to the horizontal position (attack time and duration) did not seem to affect the perceived quality of sound. With the use of sound cluster, the wavelets overlap each other, and such overlaps may preclude accurate perception of attack time and duration.

7. CONCLUSION AND FUTURE WORK

In this study, we investigated the potential of data sonification in molecular biology. The movements of wild-type and transgenic *C. elegans* were sonified using motion detection and granular synthesis, so that the sound cluster of wavelets represents the visual cluster of pixels in a perceptually matching manner. The evaluation test showed that both molecular scientists (*C. elegans* specialists) and audio engineers (non-specialists) could accurately comprehend the motion of worms through hearing, demonstrating that data sonification may have a strong potential for applications in molecular biology.

From this collaboration of biologists and audio engineers, various ideas for future directions are emerging. Technical ideas for improving sonification include the use of spatial audio, different kinds of mapping in the synthesis, and using acceleration as a parameter instead of position. However, beyond the technical ideas, we came to realize the potential of sonification in discovering new knowledge in molecular biology. Investigating the rhythmic aspects in the dynamic motion of worms would be of interest, such as the pumping gesture observed at the throat of a worm. Sonification may also be useful for observing the behaviors of model organisms.

In this project, we sonified the already-visible aspects and discovered that sounds can convey some information that we tend to dismiss with vision. However, the sonification of non-visual aspects in biology is a further promising direction. "Listening to the phenomena we cannot see at all" may lead to the most fascinating new discoveries, and seeking such model examples is a desirable next-stage goal.

Acknowledgments

We would like to thank Peter Wang for his generous support in preparing this manuscript. This work was supported by the Kawai Foundation for Sound Technology and Music, Japan.

8. REFERENCES

[1] G. Kramer *et al.*, "The Sonification Report: Status of the Field and Research Agenda," Prepared for the National Science Foundation by members of the International Community for Auditory Display Editorial Committee and Co-Authors, 1999.

[2] S. Barrass and G. Kramer, "Using sonification," in Multimedia Systems No. 7, 1999, pp. 23-31.

[3] T. Hermann *et al.*, "Vocal Sonification of Pathologic EEG Features," in Proceedings of the 12th International Conference on Auditory Display, London, 2006, pp. 158-163.

[4] F. Dombois, "Auditory Seismology on Free Oscillations, Focal Mechanisms, Explosions and Synthetic Seismograms" in Proceedings of the 2002 International Conference on Auditory Display, Kyoto, 2002, pp. 1-4.

[5] S. Brenner, "The genetics of Caenorhabditis elegans," in Genetics No. 77, 1974, 71-94.

[6] W. S. Yeo and J. Berger, "Raster Scanning: A New Approach to Image Sonification," in Proceedings of the International Computer Music Conference, 2006.

[7] K. Jo and N. Nagano, "Monalisa: See the Sound, Hear the Image" in Proceedings of International Conference on New Interfaces for Musical Expression, 2008.

[8] S. Barrass, N. Schaffert, and T. Barrass, "Probing Preferences between Six Designs of Interactive Sonifications for Recreational Sports, Health and Fitness," in Proceedings of ISon 2010, 3rd Interactive Sonification Workshop, Stockholm, 2010.

[9] J. M. Pelletier, "Sonified Motion Flow Fields as a Means of Musical Expression" in Proceedings of International Conference on New Interfaces for Musical Expression, 2008.

[10] J. E. Boyd and A. Sadikali, "Rhythmic Gait Signatures from Video without Motion Capture," in Proceedings of the16th International Conference on Auditory Display, Washington, D.C, 2010, pp 187-191.

[11] J. M. Pelletier, "Perceptually Motivated Sonification of Moving Images," in Proceedings of the International Computer Music Conference, 2009.

[12] C. Roads, Microsound, MIT press, 2002.

[13] C. C. Mello, *et al.*, "Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences," in EMBO J. No. 10, 1991, pp. 3959-3970.

[14] Caenorhabditis Genetics Center, University of Minnesota. URL: http://www.cbs.umn.edu/CGC/

[15] P. G. Okkema, *et al.*, "Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans," in Genetics No. 135, 1993, pp. 385-404.

[16] C. A. Peixoto, e*t al.*, “Ultrastructural analyses of the Caenorhabditis elegans rol-6 (su1006) mutant, which produces abnormal cuticle collagen,” in J. Parasitol No. 84, 1998, pp. 45-49.

[17] Cycling 74, Max 5. URL: http://cycling74.com/

[18] E. Zwicker and H.Fastl, Psychoacoustics: Facts and Models, Springer, 1999.

[19] S. McAdams *et al.*, “Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes,” in Psychological Research 58, 1995, pp. 177–192.