Motion Tracking of a Fish as a Novel Way to Control Electronic Music Performance

SHALTIEL ELOUL, GIL ZISSU, YEHIEL H. AMO AND NORI JACOBY

The authors have mapped the three-dimensional motion of a fish onto various electronic music performance gestures, including loops, melodies, arpeggio and DJ-like interventions. They combine an element of visualization, using an LED screen installed on the back of an aquarium, to create a link between the fish’s motion and the sonified music. This visual addition provides extra information about the fish’s role in the music, enabling the perception of versatile and developing auditory structures during the performance that extend beyond the sonification of the momentary motion of objects.

Popular electronic music performances are situations in which highly structured music defines the mood and social interactions among people within the same space. This style of music typically combines accessible musical phrases with variations that enhance the experience for the human ear [1]. Many clubs and electronic venues also create multisensory environments in an attempt to enhance club-goers’ experiences. Introducing an interactive musical aquarium is one such enhancement, creating subtle similarities between the movement of fish in an aquarium and the electronic music style.

From auditory to visual domains, fish exhibit a wide variety of swimming patterns and behaviors in response to their environments. Their movements are visually attractive to the human eye and have also proven to be valuable in cognitive studies [2] and psychological therapy [3]. The hydrodynamic shape and body movements of fish (which are related to their size and type) merge into the aquatic locomotion that has captivated human onlookers for thousands of years [2,4]. The earliest recorded reference to fish motion is perhaps that of Aristotle, who, in the 4th century BCE, tried to relate a fish’s fins to its straight motion, somewhat anticipating Isaac Newton’s identification of the fins as the physical basis of the generation of thrust for locomotion in water in the 17th century [2].

This unique locomotion can be both patterned and random, and therefore our translation of the fish’s movement into musical syntax observed the following rules: On the one hand, the fish’s repetitive paths could be captured as rhythm, while on the other hand, occasional random movements could trigger surprising musical responses and effects. An electronic club music genre that has analogous properties is minimal techno [5], characterized by slowly developing rhythms that exploit the use of repetition while making generous use of synthesized sounds and prerecorded realistic sounds. This genre is an inspiration for our work on translating the motion of a fish into various electronic music styles as a new way to express the interaction between visual and auditory senses in club and electronic music scenes.

Music computing artists have been increasingly interested in identifying new methods of expressive engagement with interactive music systems [6–8], i.e. systems that express gestures to generate and control musical parameters and signals for virtual instruments [9–13]. Another emerging area of study is the sonification of various natural phenomena by analyzing data or tracking objects. This includes water sonification [14], music based on climate [15], music maps of the routes of stars [16] and sounds based on molecules and biomolecules [17, 18]. The aim of our work is to create an expressive interactive performance with a stimulating phenomenon, in this case the sonification of fish motion, by using motion tracking.

Tracking visual motion (e.g. dancing, natural phenomena or animals’ motion [19]) in real time and sonification of motion are now relatively manageable tasks owing to the wide availability of high-speed cameras and tracking devices such as the Kinect [20, 21]. Yet the translation of these tasks into intriguing onstage musical performances remains challenging, both technologically and artistically [22]. This difficulty

©2016 ISAST  doi:10.1162/LEON_a_01072

LEONARDO, Vol. 49, No. 3, pp. 203–210, 2016 203
is likely a result of the significant gap between the demonstration of an idea in a research framework (“in the lab”) and the creation of a compelling live performance. When performing, the artist must consider many parameters, such as the quality of the music, visual interpretation, the audience's ability to understand the artistic ideas and, above all, the artist's ability to maintain the audience's interest throughout the performance. Therefore, the artist's challenge is to turn a new technological art idea into an enjoyable and pleasant experience for the audience in a live show, exhibition or club performance. Here, we present the idea of creating and playing music generated by the motion of a fish, which we have developed into an onstage performance.

**PERFORMANCE BACKGROUND**

The first major work to address the challenge of translating fish movement into music was produced at the San Francisco Tape Music Center by Ramon Sender in 1962 [23]. In his *Tropical Fish Opera*, four instrumentalists used a tank to create music scores for an improvisational performance. The players improvised the music by drawing staff lines on the fish's tank and trying to play the positions of the fish as it swam up and down in the tank.

Sonification using optical tracking was proposed by Pendse et al. in 2008 [24] and more recently by Baldan et al. in 2012 [25]; these works describe the idea of multiple fish sonifications created according to the animals' size and location in an aquarium. The studies demonstrate a restricted method of sonification, which is important for the general study of sonification and the implementation of sound from visual events. However, their auditory results did not seem to attempt to stimulate the listener's attunement to various music phrases or to develop music structures over time. Moreover, their works were focused mostly on solving technological issues in sonification, such as motion tracking with multiple objects, rather than directly applying it to the creation of music or to the combination of their technology with time-based performance.

In 2011, Nikolaidis and Weinberg [26] created a general model for dynamic sonification of visual movements into “low-level” elements such as pitch and panning as well as “high-level” elements such as melodic attraction. Still, music that relied solely on this model would be limited to tracked objects over a specific time. To avoid such specific momentary movements, we used a background screen that can interact with the movement of the tracked object, rendering the melody and musical ideas more complex and diverse. For example, “painting” the movement of the fish in real time can provide a visual record of the fish's routes or previous positions, as shown below in the section “Audio and Visual Translation Patches.” Then, after this information has been captured on the screen, we can use it to implement large-scale musical ideas (such as create rhythm, arpeggios or loops) without losing the relation between the music and the fish's motion. Above all, we must also take into account aspects of visual and music perception, aesthetic value and logistical organization, so as to provide a performance that appeals to both musicians and nonmusicians.
Sketches of the aquarium equipment and photographs from live installations are shown in Figs 1 and 2, respectively. We positioned the fish in an aquarium (Fig. 1a) onstage, as depicted in the upper photo in Fig. 2. To detect the three-dimensional motion of the fish, we placed two FireWire cameras at the top (Fig. 1b) and the right side of the aquarium (Fig. 1c). Three parameters—position, velocity and acceleration—are translated to control musical gestures and triggers. To compute the fish's motion and all music elements, we used two laptop computers (Fig. 1d), a synthesizer (Fig. 1e) and one portable audio card (Fig. 1f).

We further combined a visual interpretation with an LED screen (Fig. 1g) positioned on the back of the aquarium. This visual element is controlled by the fish's motion and by music triggers using a third laptop. We found that this added visual interpretation is crucial to the audience's understanding of the interaction between the fish and the music, as we explain below.

### APPLICATIVE METHODOLOGY

The live and time-based art we present addresses a central concern in the engagement of technology with the interactive design of musical and visual experience. The experience of the performance relies on the ability of the musicians to integrate independent fish motion into familiar structures of electronic music. These involve loops, DJ-like effects such as panning and filtering, and creating live melodic layers over an existing electronic track. To connect the audience to the performance, we decided to avoid demanding cognitive ideas, which might easily distract listeners from the integration of the visuals and the music. On the other hand, we felt that the musical concept should be continuously developed and varied to maintain the audience's curiosity and the performance structure. Therefore, we attempted to develop the performance slowly, moving from the most transparent interactions between the fish and the music to more complex interactions. As we designed the performance, we examined a number of aspects in order to understand how to maximize the interaction among the listener, the fish's motion and the music.

We swiftly realized that playing music based on more than one fish makes the connection between the sound and the motion of the fish difficult to capture and barely understandable; focused on deciphering one fish's part in the creative process, the listener becomes distracted from the whole experience. We also realized that visual elements behind the aquarium play an important role in allowing the listener to understand the fish's role. Our choices of visual effects are relatively easy to understand and constitute a "bridge" between the fish's motion and the music, making the experience of the performance comprehensible and thus allowing the listener to enjoy the performance. Furthermore, in order for the listener to distinguish between the background music and the fish's role, the musical style needs to be accessible and full of constant, easily recognizable patterns. The background music cannot involve solo lead parts, as those could confuse the listener.

Given the logistical and aesthetic aspects of the performance, we built a portable aquarium with a wheeled stand, allowing for rapid construction and portability between installations (lower photo in Fig. 2). The fish's natural behavior (e.g., swimming speed) and our ability to maintain necessary water conditions in the aquarium (e.g., temperature, oxygen bubbling and filtering) are other essential aspects of the performance. Additionally, for ethical and practical considerations, we had to ensure that the fish's life is pleasant before, during and after the show by allowing it to move in a natural way. Therefore, we designed the aquarium with a separator on the side of the tank that provides a hiding place (Fig. 1h), which is common in many fish tanks. We positioned all the aquarium equipment, including the bubbler, filter and heater, within the hiding place, allowing the fish to live in comfortable conditions. It is important to mention that our fish, a dominant red zebra cichlid, did not show any
sign of distress before, during or after the performance. It ate normally, moved calmly and explored the tank, showing some curiosity by tracking the visualization on the screen at the back of the aquarium. The fish could also move into its hiding place at any time.

TECHNOLOGY AND PERFORMANCE METHODOLOGY

We developed a modular program using a Max/MSP/Jitter [27] environment, containing three patches that allow the use of three standard laptops, each equipped with a 2.4 GHz processor. In this way, we smoothed the motion of visual effects and avoided significant digital signal latency (DSP). The first patch controls image processing for tracking the fish and translating its motion into a position, velocity (speed) and acceleration. The second patch handles music translation and DSP, and the third controls the visualization on the LED screen at the back of the aquarium.

We implemented this tracking process using the computer vision for Jitter library [28]. Two low-latency FireWire cameras are used to capture the fish at a resolution of 320:240 pixels, yielding up to 25 frames per second (fps) (Fig. 1b,c). We implement a paradigm for tracking the fish in the performance as follows: During the first step, we use normal image handling to optimize contrast and brightness and thus make the colors more distinctive and compressed. In the second step, we apply an unusual background subtraction, as normal continuous background subtraction [29] is not applicable to this system, given that the fish may move slowly or remain still, leading to an unsatisfactory detection of position. To overcome this issue, we apply a static background subtraction that is renewed on the empty half of the aquarium whenever the fish is swimming in the other half. This tactic helps to make the subtraction far enough from the fish’s position or that of its shadow. Next, blob tracking (which identifies position and size) is introduced from the cv.jit toolbox, which uses a binary image followed by a Kalman filter to track the fish’s true position [29]. In the last step, we simply use a low-pass filter of the tracked position while maintaining a practical latency—around 50 milliseconds (ms). The filter takes a series of tracked matrices and cuts off any large distance fluctuations by applying the following formula:

\[(\text{New position}) = (\text{last position}) \times 0.95 + (\text{new detected position}) \times 0.05\]

This low-pass filter is necessary for a stable, smooth estimation of the fish’s speed. The fish’s coordinates indicate its position at any frame, and its first-order discrete derivative magnitude (scalar speed, which we denote as \(V\)) is then evaluated using the following formula:

\[V = \sqrt{\frac{\Delta x^2}{f} + \frac{\Delta y^2}{f} + \frac{\Delta z^2}{f}}\]

Where \(f\) is the frame rate of the input signal, \(\Delta x\), \(\Delta y\) and \(\Delta z\) are the differences between two subsequent frames in the tracked position of \(x\) (horizontal axis), \(y\) (depth axis) and \(z\) (vertical axis and the height of the aquarium), respectively.

In the same manner, the acceleration magnitude (\(a\)) describes large changes in the fish’s motion, as follows:

\[a = \frac{\Delta V}{2f}\]

After the tracking paradigm is computed on the first laptop, the data is sent wirelessly via TCP/IP protocol to audio translation and to the visual translation patches on the two other laptops.

AUDIO AND VISUAL TRANSLATION PATCHES

The fish’s position, speed and acceleration parameters are translated to shape music gestures and triggers via four presets—creating loops, music dialogue, DJ gestures and melody composition—that control and influence the music in different ways to achieve a variety of experiences for the listener.

Creating Loops

This preset translates the fish’s movement by painting cells with respect to its \(x–z\) pathways in the aquarium. These are visualized on the screen in the background (Fig. 3). Each cell that is painted by the fish triggers a sample in a matrix as a standard loop machine. The rows of the matrix describe the sampled sound, which we prepare in advance, and the columns are the metrical position within the bar (\(\frac{1}{32}, \frac{1}{16}\) or \(\frac{1}{8}\) of a bar). We use samples from electronic drum kits, glitches and other noisy sounds. Delay and reverb effects are added when the fish arrives at corners. The loop results in an abstract structure when using 32 beats per bar or a repetitive rhythm and a danceable loop when using 8 beats per bar (Fig. 3).

In addition, we can predetermine the beats per minute (bpm) of the loop by mapping the speed of the fish to an experimental value between \(V\) (speed, \(a\rightarrow v_{\text{max}}\)) → 120–240bpm.

Music Dialogue

In this preset, the aquarium is “split” into two sides: left and right (Fig. 4). Each side is programmed for a different
Eloul et al., Motion Tracking of a Fish as a Novel Way to Control Electronic Music Performance

The fish controls the type of music and visual effects by swimming in each section of the aquarium. The visualization of the LED screen uses OpenGL in Jitter visual programming. This preset allows clarity via sharp shifts in the music according to large changes in the fish’s position, helping the listener understand the components of the system: the fish’s motion, the visualization of the screen and how the two merge into music gestures.

**DJ Gestures**

The x, y and z coordinates of the fish control musical modulation and effects: cutoff filtering, pitchbend, volume and left-to-right panning. The fish alone controls the background music, because its position and speed are mapped to those parameters. Table 1 shows how these effects are mapped according to the motion of the fish in the aquarium.

Given our goal of making gestures that are easy for a listener to understand but do not distort the background music, gesture mapping is determined by trial and error. We combine the background LED screen with visual effects designed in Max/Jitter to enhance the connection between the motion of the fish and the musical gestures. One example of a simple enhancement is shown in Fig. 5, in which a tracking ball behind the fish changes shape, size and speed vibration bits according to the fish’s velocity and amplitude peaks in the music audio (made using OpenGL in Jitter visual programming).

**Melody Composition**

Baldan et al. [25] mapped sound parameters such as pitch, panning and timbre to a large number of fish according to various physical characteristics and positions, resulting in a very rich multiplicity of random sounds. In our work, the fish is in charge of the melody, in a role similar to the player of a synthesizer. The position, speed and acceleration of the fish send MIDI messages to an external synthesizer (a Novation UltraNova). The MIDI messages trigger notes at certain accelerations. Many sound presets can be used with various attack-decay-sustain-release (ADSR) envelopes, such as lead-synth, pads or strings. MIDI messages that are controlled by the fish’s position and velocity are sent to the two wheels of the synthesizer: the pitchbend wheel and the modulation effect wheel, as illustrated in Fig. 6.

The preset is designed for the performance and is well suited to function as solo parts, above the background music. The fish is restricted to playing specific types of scales (diatonic, pentatonic, etc.), and there are specific rules for each music phrase (listed in Table 2). We determined each sound

---

**TABLE 1.** Mapping effects on the master music channel (background music).

<table>
<thead>
<tr>
<th>Gesture</th>
<th>Parameter</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitchbend</td>
<td>x (0–320 pixels)</td>
<td>−12 tone to +12 tone</td>
</tr>
<tr>
<td>Cutoff filter</td>
<td>y (0–240 pixels)</td>
<td>500 Hz–20 KHz</td>
</tr>
<tr>
<td>Panning</td>
<td>x (0–240 pixels)</td>
<td>100% left → 100% right</td>
</tr>
<tr>
<td>Volume</td>
<td>V (speed, 0–v_max)</td>
<td>−3 dB → 0 dB</td>
</tr>
</tbody>
</table>
preset using these rules. In this way, the listener can relate the motion of the fish to familiar popular music patterns that reflect the fish's characteristic motion.

As an enhancement and development of this preset that we combine in performance, we use the visualization of the LED backscreen to play arpeggios without losing the relation between the melody and the fish's motion. When the fish accelerates or makes a turn, it creates notes that are presented as 3D balls connected by wires (Fig. 7). The notes then make an arpeggio melody, whose tempo is determined by the speed of the fish. The visualization, created with OpenGL and the open-source Java library TRAER.PHYSICS [30] to make a particle system, applies forces and handles positions in real time. In addition, when the fish bumps into one of the balls, the ball sticks to the fish, and the fish virtually "grabs" and drags the ball to change and develop the arpeggio melody.

**PERFORMANCE AND MUSIC CREATION**

All the above presets are combined into two distinct performance modes. In the first mode, the fish creates the music *ab initio*, by layers (e.g. creating a loop, adding melody and then making gestures on the music that was created). Onstage, we begin the performance with one music element—the "creating loop" preset, which gives a clear visual picture of how the loop is created by the fish. After finding an interesting loop containing the base beat, we add further elements and layers. For example, we may add arpeggios or let the fish play melodies with predefined sounds. When the music becomes rich, we let the fish become the DJ by controlling effects such as filtering, panning, volume and pitchbend. Then we can repeat the set. As audience members grow familiar with the role of the fish, they typically are able to achieve a greater engagement with the musical result. In the second mode, we set the music in the background or play it alongside performers onstage, and the fish affects the background music by controlling music parameters, changing the music atmosphere or adding a melody.

We combined supplementary media, including sample video and sample sounds, which accompany the present article [31], to illustrate the main ideas of the performance and supply additional information regarding experimental performances that examine levels of interaction. Another demo provides an audio file of unedited music created by the fish during the installation. Last, a third audio file presents the result of the fish creating an arpeggio and controlling the parameters of background samples [32].

**CONCLUSION AND FUTURE WORK**

We have explained a music technology project that tracks the motion of a fish to play and control electronic music in live performance. We also propose a new understanding of the utilization of this music technology
idea, taking it out of the “lab” and onto the stage. The complexity of the interaction between the fish and the music is reduced to a set of gestures/presets, extending earlier versions suggested by previous work [24–26] in the music performance aspect. In the same manner, we have found that trying to use more than one fish can easily garble the clarity of the interaction we seek. Our solution to the problem involves the introduction of a LED screen at the back of the aquarium, allowing the listener to understand the fish’s role in the music while also inviting new ideas for creating music.

To allow more complex gestures and presets using more than one fish, interaction experiments should first be conducted with musicians and nonmusicians in order to examine their perception of the level of interaction between the fish and the music, as well as the experience as a whole. In addition, additional music presets can be examined to elevate the performance into higher levels of melodies and interaction. For example, we examined a new preset that permits a fish to play an arpeggio with the combination of more than one fish that is painted on the back LED screen, such as by making each fish responsible for one note in the arpeggio, with the notes being changed with the position of the various fish. This technique could permit melodies to develop into more interesting and complex structures without clouding the clarity of the interaction between the music and the fish’s motion, thus providing audiences high levels of music experience and interactivity that are directly influenced by the fish behavior as an individual or in a group.

Acknowledgments
This project was supported and funded by the music project syndrome. A special thanks to Eran Shlomi, Almog Kalifa and Dima Daviddovich, who participated in the installation and design of the portable aquarium and are an integral part of the project. We acknowledge Carmel Raz for her editing, proofreading and professional review of this article. Finally, we thank Liv Fiertag for her proofreading and language advice.

References and Notes

1 A. Bennett, Popular Music and Youth Culture: Music, Identity and Place (New York: Palgrave Macmillan, 2000).
ANNOUNCING

Leonardo Art Science Evening Rendezvous (LASER)

Since 2008, the Leonardo Art Science Evening Rendezvous (LASER) series of lectures and presentations on art, science and technology has provided spaces for progressive thought leaders to come together to form community and explore the intersections of disciplinary thinking. Owing to its success and popularity, LASER has expanded beyond its birthplace in the San Francisco Bay Area, first to the U.S. East Coast, then across the Atlantic to London—the home of the first European LASER—and today continues to expand to new locations nationally and internationally. We thank all of those who have spoken at, participated in or attended LASER events throughout the years. We owe a special thank you to Piero Scaruffi, LASER founder and chair, for his inspiration and continued dedication, and to the growing list of LASER hosts around the world. To follow LASER events, see <www.leonardo.info/isast/laser.html>.
Leonardo Book Series

EDITOR IN CHIEF: Sean Cubitt
EDITORIAL ADVISORY BOARD: Annick Bureaud, Laura U. Marks, Anna Munster, Michael Punt, Sundar Sarukkai, Eugene Thacker
EDITORIAL CONSULTANT: Joel Slayton

The arts, sciences and technology are experiencing a period of profound change. Explosive challenges to the institutions and practices of engineering, art-making and scientific research raise urgent questions of ethics, craft and care for the planet and its inhabitants. Unforeseen forms of beauty and understanding are possible, but so too are unexpected risks and threats. The Leonardo Book Series, published by The MIT Press, aims to consider these opportunities, changes and challenges in books that are both timely and of enduring value.

Proposals that address these challenges in terms of theory, research and practice, education, historical scholarship, discipline summaries and experimental texts will be considered. Single-authored books are particularly encouraged.


Inquiries and proposals should be submitted to both:

Leonardo Book Series
c/o Leonardo
1440 Broadway, Suite 422
Oakland, CA 94612
U.S.A.
Email: <leonardobooks@mitpress.mit.edu>.

Doug Sery
MIT Press Books
One Rogers Street
Cambridge, MA 02142
U.S.A.

RECENT TITLES:
Laura U. Marks: Hanan al-Cinema: Affections for the Moving Image
Joasia Krysa and Jussi Parikka, eds.: Writing and Unwriting (Media) Art History: Erkki Kurenniemi in 2048
Seb Franklin: Control: Digitality as Cultural Logic
Gloria Sutton: The Experience Machine: Stan VanDerBeek’s Movie-Drome and Expanded Cinema

To order Leonardo Books, visit <leonardo.info>.