Shedding Light: Integrating Bioimaging Technologies into the Design of an Interactive Museum Exhibit

Joyce Ma
Exploratorium
San Francisco, CA USA
jma@exploratorium.edu

Eyal Shahar
Exploratorium
San Francisco, CA USA
eshahar@exploratorium.edu

Kevin Eliceiri
University of Wisconsin
Madison, WI USA
eliceiri@wisc.edu

Guneet Mehta
University of Wisconsin
Madison, WI USA
mehta.guneet@gmail.com

Kristina Yu
Exploratorium
San Francisco, CA USA
kyu@exploratorium.edu

ABSTRACT
In the past decades the field of bioimaging has experienced an explosion of technologies that have revolutionized the study of the microscopic world. Through iterative prototyping and evaluation with museum visitors, we shed light on how these technologies, specifically image recognition techniques, can be incorporated into an interactive museum exhibit to help the visiting public gain insights and make sense of dynamic, live specimens as they use a research-grade microscope. Our work indicates that the technology needs to be carefully positioned as an *aid to observation* rather than an infallible source of expertise in order to support and sustain visitors' explorations. This paper contributes to our understanding of design considerations in integrating emerging imaging technologies and techniques into museum exhibits to enrich visitors' interactive learning experiences.

Author Keywords
Image recognition; museum; informal science learning; interactive exhibits; bioimaging; microscopy.

CSS Concepts
• **Applied computing~Interactive learning environments** • Human-centered computing~Interaction design

INTRODUCTION
The scientific endeavor is increasingly characterized by the use of sophisticated instrumentation and technologies. Nowhere is this more true than in microscopy, where the advent of live cell imaging techniques used to observe living cells under a microscope, have revolutionized biology [6],[16],[19]. Using a video camera, and later, computers and digital cameras, scientists have gained extraordinary insight into the dynamic processes of life at the microscopic scale. Automated software-controlled microscopes, imaging algorithms, ultra-sensitive digital cameras, and other electronics have resulted in an explosion of advanced computational techniques that continually push the envelope of what is biologically visible [10]. They have allowed researchers to observe biological processes as they unfold and have been instrumental in making key scientific discoveries from revealing the cellular composition of the cerebral cortex to better understanding the origins of cancer.

DIS '19, June 23–28, 2019, San Diego, CA, USA

ACM ISBN 978-1-4503-5850-7/19/06…$15.00
https://doi.org/10.1145/3322276.3322299

These advances offer informal science learning venues, such as science museums and centers, natural history museums, and aquaria, exciting opportunities to open up a dynamic new world of the living microcosm for their visitors, using authentic tools and the same living specimens used by scientists. These imaging technologies also provide exhibit developers tools they may be able to adapt and integrate to better support visitors in exploring and making sense of the microscopic world.

This paper presents a case study of our attempts to adapt and integrate bioimaging techniques into an interactive, research-grade microscope exhibit at a science museum to support visitors in exploring a dynamic, living sample that such instruments now allow. More specifically, we studied three approaches in situating image recognition technology in the visitor experience:

(1) *Technology as content generator*. Bioimaging technology is used to determine what is currently under the microscope and, in turn, what content information is presented to visitors.
(2) *Technology as finder*. Image recognition algorithms are used to find, identify, and point out areas of biological interest.
(3) *Technology as tool*. Visitors are given control of an image processing module as an observation tool that they could use in their visual explorations.

The three approaches were explored in succession as cycles of an iterative development and evaluation process, with each iteration's lessons learned informing the next. The rationale for each successive approach was therefore emergent, grounded more in the constraints revealed in the prior iteration than in overarching theories of technology integration, which are still scant for bioimaging technology in the museum context. For each approach, we considered its implementation, its technical feasibility, and the viable prototype's use by museum visitors in order to surface design considerations regarding how the technology may be integrated into the museum visitors' experience.

We hope that the design lessons learned from this case study can contribute to our understanding of how bioimaging techniques and, more broadly, imaging techniques, can be used to help informal science learners explore visually complex, dynamic imagery that increasingly characterize current biology and other scientific fields that rely on detailed observations of real phenomena with sophisticated instrumentation.

RELATED WORK

The Museum Environment

Museums are free-choice, informal learning environments whose visitors are largely motivated and guided by personal interests [12],[23]. These institutions can attract a diverse audience, broad in interests, experience, and knowledge. Exhibits developed, interpreted, and otherwise curated by a museum are the primary means used to structure the visitor experience. The exhibit experience has often been characterized as unmediated and short [23]. That is, many science exhibits are designed as stand-alones, to be used without staff facilitation. Instead, physical affordances and labels are typically the only means for mediating interaction and interpretation for visitors, many of whom may be unfamiliar with the underlying scientific content. Because an exhibit experience rarely sits within a larger curriculum or a predetermined sequence within a gallery or exhibition, they are often designed as self-contained experiences that assume few pre-requisite skills. Furthermore, each exhibit experience tends to be short, measured in minutes or, more typically, seconds. For example, in the life sciences area of the Exploratorium, a science museum in San Francisco, visitors spend from a little over ten seconds to under two minutes at one exhibit [15]. The short holding time provides little opportunity to 'train' visitors on sophisticated skills.

The challenge of designing any successful exhibit, therefore, lies in quickly initiating visitor engagement in what may be an unfamiliar subject and sustaining their interest, all without staff facilitation and with limited time. A key way in which museums, especially science museums and centers, have tried to engage their visitors is through the design of interactive exhibits [13],[14],[23]. Studies have shown that interactive exhibits can foster engagement with natural phenomena, increase social interaction with other visitors, and promote content understanding and recall [2],[13],[23],[26]. In life-sciences exhibits, with few exceptions (e.g., [18]), interactivity with *living* specimens is largely realized through observational instruments, such as webcams [7], binoculars, and microscopes [11],[27], which visitors control to look at and look for organisms. Accompanying the instrument are interpretative labels, often static, that try to help visitors make sense of what they may see while encouraging their explorations.

Live Microscopic Specimens in Exhibits

Although microscopes have long been part of the visitor experience in museums, there are still few instances where the public are provided access to research-grade microscopes that they can control to explore the living microcosm. Micropia in Amsterdam, the New York Hall of Science, the Science Museum of Minnesota, the California Science Center, and San Francisco's Exploratorium are some of the handful of museums that have had long-standing installations of microscope exhibits featuring live samples.

Despite the costs associated with providing visitor access to live organisms, exhibits featuring living specimens are important parts of a life sciences gallery. Live specimens hold special appeal for visitors. The summative evaluation for the Exploratorium's Microscope Imaging Station project found a large majority (85%) of visitors felt that seeing something alive and moving made the exhibits worthwhile [11]. Similarly, in a study comparing visitors' engagement and learning at different versions of a microscope exhibit, Allen [1] found that visitors who used an interactive microscope with live specimens were more engaged and recalled more about what they saw than visitors who watched a pre-recorded video of the same specimen.

Machine Vision in Museums

In their 2017 review, Majd and Safabakhsh [22] described a variety of machine learning applications used in museums, including for collection management, data analysis, marketing, and museum planning. According to their survey, developers have also attempted to incorporate machine vision to enhance the visitor experience, generally in the form of automatic mobile guides. In these systems, image recognition modules are used to help determine what visitors may be near or attending to and bring up the content associated with that museum artifact. That is, they generate the dynamic interpretative labels for exhibits.

The papers surveyed as part of their review focused on the technical aspects of using machine learning but not the resulting visitor experience. Only one considered visitors' preferences and use patterns with a mobile application, which employed image recognition versus QR codes to identify the artwork viewed to access and display background information about the subject [33]. Work on integrating live cell imaging techniques to help museum visitors explore the dynamic images that characterize live microscopy is only just beginning.

DESIGN

Considerations

Bioimaging techniques encompass a vast array of technologies, instrumentation, and protocols, ranging from confocal laser scanning microscopy to high throughput imaging of cells used for drug discovery. A focus of the field is on the imaging of living samples, which change, move, and react in space and time. Developed in pursuit of scientific research agendas, many techniques are not readily extensible to museum exhibits. In the following, we describe characteristics that we hypothesize make for a promising technique to help visitors make sense of dynamic biological imagery.

Real-Time Results

Whereas research images can be captured for post hoc processing in a laboratory, visitors are usually watching a live video stream of a living, fluidly moving specimen. To maintain the visitor's attention requires that any bioimaging technique needs to work on the live image and output a result in well under a second. In the free-choice learning environment of a museum, a frustrated or confused visitor can readily move to another, more exciting or satisfying exhibit. Whatever image processing and recognition algorithms are to be incorporated into an exhibit, they should have short and predictable execution time, and preferably have results ready even before visitors require them. Various techniques can be applied to achieve this. For instance, images and videos can be down-sampled in pixel resolution and frame rate, and algorithms that require time to process, whether it is for accumulating video frames or performing classification, can constantly run in the background or on dedicated machines.

Aesthetics

Museum goers respond to the aesthetic quality of images displayed [32]. High-resolution, color images are important in attracting and engaging visitors. Biomarkers such as Green Fluorescent Protein (GFP), which light up when exposed to specific wavelengths of light, need to be carefully selected to be visually striking. To maintain image quality, full-color, high definition video cameras that can stream in real time are preferred to show images from the microscope.

Apprehendability

What is meaningful in research is defined by disciplinary knowledge, foundational theories, and the overarching questions that anchor and motivate the scientific effort. Museum visitors, however, come to an exhibit with little to no awareness of this larger scientific context. Consequently, the enabling techniques need to help visitors make sense of, connect with, and otherwise see meaning in what they see quickly. Allen refers to this quality as the 'immediate apprehendability' of an exhibit [1]. For a microscope exhibit, this entails highlighting what is immediately visible and recognizable to visitors. Prior evaluation has shown that visitors tend to focus on specimens that have identifiable features while passing over those that are unfamiliar [20]. Techniques that focus on long-term processes (e.g., cell division, which can take hours) or unfamiliar parts (e.g., developing sea urchin embryos, which though visually striking may also appear completely foreign and unrelatable) would, therefore, be poor candidates to integrate.

Accuracy

Some visitors may come to an exhibit with no prior experience using a microscope. While research images are often optimized to allow scientific visual interrogations, an image that a novice visitor brings into view may be out-of-focus or off-center. Techniques that assume optimal image quality would be difficult to integrate into an exhibit without making allowances for possible inaccuracies.

Image recognition techniques can be roughly split to those using machine learning and those based on signal processing. When using machine learning techniques, such as convolutional neural networks, preparing the training set requires the inclusion of a very large amount of examples of any state that the visitor may encounter. This may increase the size of the training set significantly and make it harder for the machine to distinguish subjects of interest. Heuristic, signal processing based methods avoid some of these difficulties but come with their own challenges, since the geometries they seek, such as distinct lines, complete and well-defined contours, and clear boundaries between color areas, are not always available in a suboptimal, visitor-generated image.

The above considerations led us to focus our initial work on signal processing based, heuristic image recognition techniques, although machine learning techniques are likely to give more consistently accurate results for more general applications and a wider range of specimens, and warrants future investigation. Furthermore, the specimen used in our exhibit had simple, well-defined geometries that lend itself to signal processing methods. Although our work depended on identifying, procuring, and maintaining live organisms with brilliant and robust biomarkers, we chose to focus this study on image recognition as opposed to bio-specific techniques (e.g., tracking specific GFP engineered into a transgenetic organism), which are particular to a specimen. Lessons learned from integrating image recognition techniques, we hope, would find wider applicability.

Exhibit Platform

The Exploratorium's Microscope Imaging Station facility, developed with support from the National Institutes of Health, the David and Lucile Packard Foundation, and the Troy and Leslie Daniels Fund for Life Sciences, provided the platform for this work. The Microscope Imaging Station sits along a corridor in the Living Systems gallery of the Exploratorium, a science museum in San Francisco, California with over 850,000 visitors annually. It has been used as the technological platform for a series of interactive microscope exhibits featuring live samples used in biomedical research such as dividing cancer cells, developing stem cells, crawling amoebas, and pulsating human heart cells [5],[17].

The Microscope Imaging Station has two main subsystems: a microscope backend, placed inside the museum's biology laboratory (Figure 1a) and visible to the visitor through a window, and a user kiosk (Figure 1b), where the interactions between visitors and the exhibit take place.

Figure 1. The Microscope Imaging Station, (left) the equipment inside the biology lab, and (right) the user kiosk.

Microscope Backend
At the heart of the microscope backend is a research-grade, automated Zeiss Axiovert 200M microscope. The microscope is controlled by a Zeiss MCU28 controller, which takes instructions through its serial port to control the microscope's stage (x and y position), focus, objective (lens) selection, and light choice. The instructions are generated by a desktop computer running Go2Scope, a client/server microscopy system based on the Micro-manager platform [9]. This custom software allows other computers to send instructions to the microscope over the internet or a local network.

User Kiosk
Visitors interact with the exhibit via a kiosk, using a joystick and a dial to control the microscope's stage and focus, respectively. An Epiphan DVI2USB video grabber installed on the kiosk computer captures a live video stream from a microscope camera, which is then displayed on a touchscreen. The kiosk software, written in JavaScript, overlays on top of the live video stream various graphical elements, such as buttons to control the microscope's objective and lighting, onscreen text to help visitors use the microscope and interpret what they see, and an observation toolbox that visitors can access to help see different aspects of the specimen.

The Test Sample
The microscope exhibit featured *Danio rerio*, more commonly known as zebrafish. With its clear eggshell, the early embryos of zebrafish have been essential in understanding the development of vertebrate animals, particularly the processes that control development and organogenesis [25], and have become a powerful model organism used by researchers around the world [4]. Zebrafish have recognizable and dynamic structures such as a beating heart, eyes, and blood familiar to most visitors [21] and thus was a promising candidate to feature in a microscope exhibit.

Target Audience
This study's target audience was visitors in family or peer groups, eight years and older. Most visitors who come to a science museum visit with others [3],[29], and most of the exhibits developed at the Exploratorium are designed to accommodate multiple users across a broad range of ages. Furthermore, by age ten, most youths have had some experience with microscopes [24], and with a user-friendly interface, the prototype used in this study could accommodate slightly younger children.

TECHNOLOGY AS CONTENT GENERATOR
In this approach, we attempted to use image recognition to determine what interpretative supports to display. Signal processing approaches were selected for their ease of use, relative accuracy for the zebrafish specimen and speed in deployment in the live view. The *technology as content generator* approach is most similar to prior uses of imaging technologies in museums, to bring up content for a particular artifact or artwork.

Image Recognition Module
Our intent was to access and display content that is specific to two different stages of zebrafish embryo development: 1) early, for which the content would be focused on the embryo's yolk and body, and 2) mature, for which the content would discuss the embryo's developing heart and circulatory system. (See Table 1.) Note that a zebrafish embryo can develop from the early to mature stage in less than a day.

In order for the exhibit to present content that is relevant to what is immediately visible to visitors, we adapted and tested a series of computer vision algorithms. Our selected approach recognized an embryo by the round shape of its egg shell. A mature embryo would be distinguished from younger embryos by its dark eyes. Fry, fish that have recently hatched, would also be detected by the eyes detector but would not trigger the egg shell detector.

The egg shell detector and the eyes detector were implemented as two plugins for ImageJ, a Java based image processing platform [30], and made accessible to the kiosk software by running as an HTTP server on the microscope backend computer. The round circles of the eggs were identified using the Circle Hough Transform [8], and the eyes, characterized as dark blobs, were recognized using ImageJ's built-in particle analyzer. Table 1 summarizes the expected responses from the two detectors when presented various images.

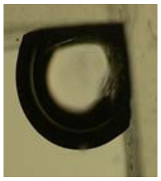
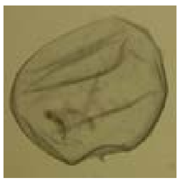
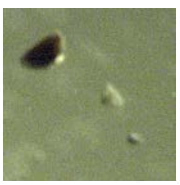
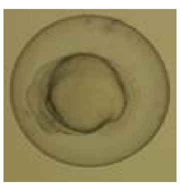
Case	Example Image	Egg Shell Detector	Eyes Detector
Nothing		Negative	Negative
Air Bubble		Negative	Negative
Empty Egg Shell		Negative	Negative
Debris		Negative	Negative
Early Embryo		Positive	Negative
Mature Embryo		Positive	Positive
Fry		Negative	Positive

Table 1. Expected detectors' outcome.

Technical Evaluation

A successful implementation of the algorithm meant not only a high rate of True Positives (TP) for each detector, but also a low rate of False Negatives (FN). Moreover, what matters to the visitor is whether or not the algorithm identified the entire content of the slide correctly.

For evaluation, a set of 101 images containing images of embryos at different stages as well as fry and empty shells (residue from recently hatched fish), was presented to the algorithm for analysis. Images containing bubbles or debris and empty slides were not included in the evaluation.

The egg shell and eyes detection algorithms were accurate at 69% and 77%, respectively, for the 101 test images used. (Accuracy is defined as the sum of TPs and True Negatives (TNs) divided by the total count.) Taken together, the accuracy of distinguishing a fry, a mature or young embryo, and no specimen was 49%. Most FNs occurred for images not fully focused but in which visitors could still clearly identify the round shape of an egg shell and a dark blob of a mature embryo's or fry's eye. Other errors took place when the egg detector misidentified an empty egg shell that still appeared round as if it contained an embryo, or when the eyes detector falsely identified dark areas in the fry's tails.

Embryos that were clipped on the edge of the image often went undetected by the algorithm. Had we treated these as TNs then the accuracy of the egg detector would have risen to 84%, and accuracy of distinguishing the sample to 62%. We, however, used the more conservative calculations because they better reflected what humans would have considered as a specimen being in view.

We did not consider a 49% accuracy rate high enough to drive content selection. If implemented, the image recognition would have pulled up incorrect content that had little to do with what a visitor was viewing about half the time.

Lessons Learned

The best solution may not depend on image recognition algorithms. Throughout our work, we were constantly reminded of how much better humans are than machines in identifying objects under suboptimal conditions. Although we could have continued to refine the algorithms in an attempt to increase the accuracy, we found other means to identify what embryo was in view. These methods relied on finding gas permeable material in which we etched micro-chambers, small enough to hold a single embryo each. Because of the highly predictable progression of embryo development, by tracking time and the position of each chamber, the system would know with 100% accuracy both the location of each embryo and its developmental stage, and, hence, the features (e.g., heart, eyes) that may be visible. This redesign points to an important lesson learned: machine vision may not necessarily be the best or only solution. This is particularly the case when designing with living organisms. A viable solution can emerge through coordination among wetware, materials and procedures for maintaining and displaying a specimen, hardware, and software.

TECHNOLOGY AS FINDER

Given that other means were found to determine the zebrafish features that may be visible, we sought to explore another potential use for image recognition in supporting visitors' exploration of the live specimen. Earlier evaluation found that visitors were interested in the heart [21], but it was unclear if visitors could readily find it. Therefore, in this second approach, we experimented with integrating imaging technologies to find and point out interesting and familiar features (i.e., the heart) for visitors.

In this *technology as finder* approach, a *Motion Detector* was incorporated into the exhibit interface. Onscreen prompts pointed to this aid and encouraged visitors to use it if they needed help finding rhythmic movement, such as a heartbeat. When visitors touched an onscreen button labeled 'detect', the *Motion Detector* would overlay a small blue bounding box on the live video image around areas where periodic motion was detected (Figure 2). Once introduced, the *Motion Detector* was accessible to visitors at any time by pressing an onscreen button. In addition to detecting beating hearts, it also found regions of pulsating blood flow, another area with periodic movement. However, the onscreen text introduced the *Motion Detector* as a means to help visitors find the heart.

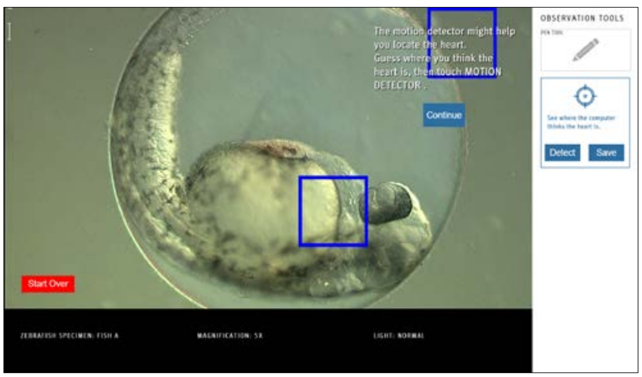


Figure 2. Technology as finder. The *Motion Detector* indicates the exhibit's best guess at where the embryo's heart is. In this screenshot, it shows two candidates for the heart location, one correct and the other (off the embryo), incorrect.

Reflecting on the previous iteration we realized that no matter how much we improved our image recognition technology there would always be the possibility of errors, and when these are framed as absolute facts they may cause visitor confusion and frustration. Therefore, we sought to position the image recognition as a fallible technology that did not always have the right answers. Instead, what it detected was presented as best guesses. For example, the use prompt was framed as, "See where the computer thinks the heart is." Visitors were then encouraged to compare where they thought the heart was to the computer's location. This is in contrast to our first approach that used image recognition to generate the accompanying content information assuming it correctly determined what specimen was in view.

Image Recognition Module

Development of the image recognition module focused on detecting periodic motion within the live video image that might indicate the location of the embryo's beating heart. In order to detect the heart's location, every frame captured by the algorithm was converted into a black and white image and then decimated. The decimation served both as a way to reduce the amount of computation needed as well as a noise reduction mechanism. The result was subtracted from the previous decimated frame, and an absolute value, or score, was taken. This created a new image where brighter pixels signify greater change between frames. This result was then fed to a first order, low pass filter to eliminate noise and abrupt movements and keep slow, repetitive changes. See Figure 3 for a flowchart of this process.

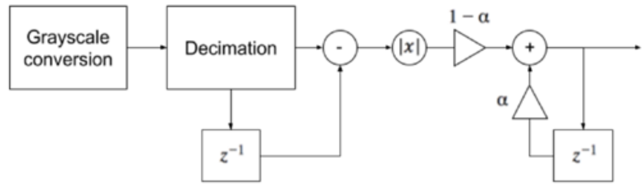


Figure 3. Heart detection algorithm flowchart.

Up to three locations of motion were presented to the visitor. To select these locations, the algorithm found the pixel with the highest score and stored it as a candidate. Then, it eliminated the pixels within a given radius around that highly scored pixel and repeated this process two more times to identify the second and third candidates.

To determine which of the three final candidates would be presented on the kiosk interface, each candidate's score had to be above a predetermined threshold. The second and third candidates' scores also had to be within a given ratio to the score of the first candidate. This way, the second and third candidates would not display if the first candidate scored significantly higher.

Technical Evaluation

The performance of the algorithm was evaluated with 147 snapshots taken by visitors as they used the *Motion Detector* over four days to allow us to assess a variety of specimens and images as captured through actual use. In our accuracy calculations, if any of the up to three candidates appeared as blue bounding boxes over an embryo's heart, that was counted as a TP. A TN occurred when no box appeared and there was no heart visible. Given these criteria, the algorithm had an accuracy of 74%.

Visitor Evaluation

Two separate evaluation studies were conducted on different days with museum visitors to assess the *technology as finder* approach. The first used interviews administered immediately after exhibit use to gauge visitors' perceptions of the usefulness of the *Motion Detector*. The second study looked at visitors' naturalistic behavior at the exhibit to ascertain if visitors used the finder successfully to locate the heart. The findings serve to identify the promises and pitfalls of the *technology as finder* approach to the integration of image recognition into the microscope exhibit.

Interview Study

Method. An evaluator approached every other visitor who crossed a predetermined imaginary line near the exhibit and who appeared to be in a group. She asked if the group would be willing to use an exhibit and answer a few questions immediately afterwards. In keeping with the target audience for the exhibit, we only recruited visiting groups where every member appeared 8 years or older. We further chose to limit the group size to two or three because the exhibit was designed to comfortably accommodate at most three people at a time.

Visitors who consented to the study were asked to use the exhibit however they liked, and to let the evaluator know when they were done. Immediately after their exhibit experience, the evaluator asked one randomly selected person in the group a series of questions. If the group had not used the *Motion Detector*, the evaluator asked the selected visitor to use it during the interview.

In total, 33 visitors used the exhibit and completed their interviews. The demographic composition of the study participants and the type of group they were in are shown in Table 2.

Demographics		***Technology as finder* approach**	***Technology as tool* approach**
Age	Adult	21	21
	Teenager	9	7
	Child	3	5
Gen-der	Female	17	16
	Male	16	17
Group Type	Adult(s) with Minor(s)	20	12
	Adults	12	17
	Minors	1	4
Total		33	33

Table 2. Demographic composition of visitor interviewed.

Results. A large majority (77%) of the visitors expected the *Motion Detector* to help them find the heart. For those visitors who reported finding the heart easily (76%), the finder then became superfluous: "*I feel like I could find the heart on my own. So, it wasn't as useful since I could already see it moving* (V18)," and "*it just showed me what we had already found* (V8)."

For seven visitors, the *Motion Detector* gave confusing results because they either expected it to clearly identify the heart: "*It has a few rectangles which can be confusing* (V9)." Or, it misidentified its location: "*It boxed a region that was not the heart* (V13)," and "*the box didn't appear over the spot we circled [with the heart]* (V6)." We note that all but one visitor reported eventually finding the heart.

Despite the difficulties, the evaluation found a large majority (88%) of the visitors interviewed thought the technology was useful in some regard. For visitors who did not readily see the heart (24%), the *Motion Detector* helped them find it: "*It helped me to track down the heart* (V7)," and "*[it was] helpful for [you] if you haven't seen zebrafish embryo or if you don't know anything about anatomy* (V10)."

Even visitors who found the heart readily thought the finder bolstered their confidence; 79% of the study participants appreciated the confirmation the Motion Detector provided: "*[It was useful] just to confirm that we were on the heart and looking where we were supposed to be* (V15)," and "*it makes you feel like you're correct. People like to feel like they're right* (V19)."

Naturalistic Observation Study

Method. We videotaped visitors' naturalistic behavior at the exhibit, collecting about 18 hours of video of what visitors saw under the microscope and their interactions with each other and the exhibit. (See Figure 4.) Museum staff and evaluators were not present as visitors used the exhibit. A program logged when visitors opened and closed the *Motion Detector*.

Figure 4. The video captured and used in naturalistic observation study. This includes a screen capture of the user kiosk (upper left) and video of visitors at the exhibit.

Using the videos, a human coder systematically sampled for observation every third visitor who appeared at least 8 years old and who came with a group. In the subsequent analyses we also eliminated visitors who stayed at the exhibit for less than 5 seconds, which was not an adequate amount of time for them to engage in any meaningful exhibit use. For each visitor who met our criteria, we noted when she or he stopped at and left the exhibit, and coordinated that timing information with the computer logs.

Subjects. The demographic information of the visitors systematically selected for observation is shown in Table 3.

Results. We found that 42% (64/154) of the visitors used the *Motion Detector*. For 53 of these 64 (83%) visitors, the finder successfully found the heart at some point during the visitor's time at the exhibit. 'Success' was assessed by a human coder who looked at the blue bounding boxes, which indicated where the algorithm detected rhythmic movement, and determined if the heart was in a box.

Demographics		Technology as finder approach	Technology as tool approach
Age	Adult	82	65
	Minor	72	48
Gender	Female	73	65
	Male	81	48
Group Type	Adult(s) with Minor(s)	121	78
	Adults	22	21
	Minors	11	14
Total		154	113

Table 3. Demographic composition of subjects selected in naturalistic observation study.

To determine if the *Motion Detector* may have helped sustain engagement, we compared the holding time for visitors who successfully used the finder versus for those whom the *Motion Detector* failed to find the heart. Exhibit holding time, the time spent at an exhibit, is a well-established metric for visitor engagement in museum studies [23],[31]. Because the holding times failed the Shapiro-Wilk Test ($p < 0.001$) for normal distribution, we performed a Box-Cox transformation, with $\lambda = 0.03$, and conducted all subsequent analyses on the resulting statistically normal dataset. An independent t-test found no significant difference in holding time between the 53 visitors for whom the finder found the heart (reverse-transformed M = 2:51, SD = 0.91) versus the 11 visitors for whom the finder failed to locate the heart (reverse-transformed M = 1:36, SD = 0.91), $t(13) = 2.07$, $p = .058$. However, we note that the small number of visitors who failed to find the heart with the *Motion Detector* made it difficult to detect any difference statistically.

Lessons Learned

These results point to several challenges and promises in the *technology as finder* approach.

An exhibit cannot depend solely on image recognition to help disoriented visitors. Any image recognition algorithm is only as good as the live feed from the microscope, and that relies on visitors' bringing the object of interest into focused view themselves. Image recognition, therefore, would have limited utility for the visitor who is at a complete loss for what to look for and how to control the microscope to find and look for it. Other means to first ensure a clear image (e.g., setting limits on the focus and stage positions) are critical to providing a meaningful exhibit experience.

Confirmation can be an important role for image recognition technology. Alternatively, for visitors who can readily find the object of interest, image recognition may appear superfluous. It may even dilute a few visitors' sense of agency in driving their own visual explorations, and thus compromise engagement. According to one visitor, "*I just enjoyed finding it [the heart] on my own* (V4)." Yet, even for these visitors, the technology found utility in validating their observations. It can be used to give visitors some additional confidence in what they see and do with sophisticated scientific instrumentation.

Visitors can tolerate less than perfect image recognition results when they are appropriately framed. Although a few visitors interviewed reported feeling confused or frustrated by its inaccurate or imprecise results (i.e., highlighting regions of rhythmic motion instead of simply the heart), an imperfect algorithm did not derail these visitors' exhibit experiences. Framing the output as 'best guesses' versus 'answers' may have helped in this regard. In fact, in the observation data, we saw some visitors repeatedly trying the *Motion Detector* when the initial attempt failed, at times refocusing or otherwise improving the image until the finder located the heart. In addition, giving visitors multiple, alternative ways (e.g., changing the light source) to visually explore the specimen also helped mitigate the need for perfect results. As one visitor explained, "*It's always fun to try and see it [the heart] and see how it's beating with the different controls* (V61)."

TECHNOLOGY AS TOOL

Although the *technology as finder* approach showed promise, we wondered if giving visitors more control over the technology would improve their experience (i.e., increase holding time and positive self-reports). In this third, *technology as tool* approach, the image recognition module was re-framed as a *tool* that visitors could control to search for the heart and other pulsing areas. Instead of the technology finding and pointing out these areas of interest, this version tried to encourage visitors to make conjectures about the heart's location and to use the tool to look for not only the heart but also other regions that move with the heartbeat (i.e. blood cells that are visible and pulsate in time with the beating heart). Prior work suggests that more open-ended experiences correlate with higher visitor engagement at exhibits [28].

To do so, the former *Motion Detector* was redesigned as a click-to-activate tool and renamed the *Beat Detector*. When activated, a circle, or lens, appeared that visitors could move over areas of interest in the live video feed from the microscope. If the module detected rhythmic motion within the circle, its rim would change colors from red (no beating detected) to green (beating detected) as shown in Figure 5. Compared to the *technology as finder* design, this interface necessitated visitors to actively note where they think the heart and other rhythmic motion may be, by moving the tool to the location they suspected. In addition, the onscreen text encouraged visitors to look for rhythmic motion in areas other than the heart. Other aspects of the prototype were also changed to improve usability (e.g., eliminating under-utilized interface features).

Figure 5. Technology as tool. The *Beat Detector* can be dragged over the live image to detect rhythmic movement. Green denotes that the system has detected rhythmic motion within the circle. In this screenshot, the *Beat Detector* has been dragged over the heart and shows green.

Image Recognition Module

The image recognition algorithm was identical to the one used in the *technology as finder* approach. However, instead of sending results from the image recognition algorithm to the kiosk upon visitor input, in this approach the algorithm constantly reported its current finding, allowing the tool's color to change immediately in response to its location as visitors moved it on the screen.

Visitor Evaluation

Interview Study

Method. The method used was the same as for the evaluation conducted for the *technology as finder* approach.

Participants. The demographics composition of the 33 study participants is shown in Table 2. Chi-square tests were performed to ascertain the comparability of the study participants across the two approaches, and found no significant difference in the interviewees' age, gender, or group composition: χ^2 (2, N = 66) = 0.75, p = 0.69, χ^2 (1, N = 66) = 0.06, p = 0.81, and χ^2 (2, N = 66) = 4.66, p = 0.10, respectively.

Results. Visitors' perceived usefulness of image recognition was very similar when it was positioned as a tool versus finder. Quantitatively, a chi-square test found no statistically significant difference in the percentage of visitors who reported any benefits in the two approaches, χ^2 (1, N = 66) = 2.39, p = 0.12. Qualitatively, visitors described similar benefits: Especially for those who did not readily see the heart, visitors talked about how the *Beat Detector* tool helped them locate the heart: "*[It] tracks where beating is, and it'd be like a hot and cold game. It's here, or it's not here* (V36)." And, "*it's easier when this highlights it [heart]* (V42)." Even for visitors who found the heart without the tool, some found the confirmation helpful: "*When I try in different parts, then I can [be] sure this is the correct place* (V45)."

Similarly, reframing the technology as a tool did not seem to mitigate visitors' confusion and frustration. A chi-square test found no statistically significant difference in the percentage of visitors who reported some type of dissatisfaction with the output of the image recognition module when it was incorporated as a tool versus as a finder, χ^2 (1, N = 66) = 1.22, p = 0.27. Qualitatively, their complaints were similar with one notable exception. Again, a few visitors struggled with the inaccuracy of the image recognition module: "*It's a little bit confusing, like I can see blood pulsing here, but it says there isn't anything there*" (V39). But, this time a few participants also talked about wanting specifically a heart detector similar to the *Motion Detector*, the finder included in the prior version: "*It just tells me that there's beating motion, not the confirmation that I found [the heart]* (V43)."

Despite these complaints, all the visitors interviewed were able to find and observe the heart. This speaks again to the importance of not depending solely on image recognition modules to help visitors have a meaningful exhibit experience.

Naturalistic Observation Study

Method. We followed a similar data collection and analysis procedure as that used in the observation study for the *technology as finder* approach. In this case, we logged when visitors opened and closed the *Beat Detector* tool, and determined when the tool detected rhythmic motion inside the circle.

Subjects. The demographic information of the visitors, as determined by appearance, is given in Table 3.

Results. The analyses found that 30% (34/113) of the visitors who stopped at the exhibit used the *Beat Detector*. A chi-square test did not find any statistically significant difference in the percentage of users between the *technology as finder* versus *tool* approach, χ^2 (1, N = 267) = 3.69, p = 0.055.

Of these 34 visitors, 15 (44%) successfully used the *Beat Detector* to find the heart. "Successful" use means that at any time during exhibit use, the tool turned green and stayed green for more than 2 seconds while positioned over the heart. A human coder determined if the zebrafish heart was inside the green circle. Compared to the *technology as finder* approach, when the image recognition algorithm was integrated as a tool, visitors were *less* successful at finding the object of interest, χ^2 (1, N = 98) = 13.88, p < 0.001.

To compare these visitors' holding times, we used a Cox-Box transformation with λ = 0.074 to normalize the data, which was right-skewed. An independent t-test found that the 15 visitors who managed to use the tool to find the heart stayed longer at the exhibit (reverse-transformed M = 2:38, SD = 0.97) compared to the 19 who could not (reverse-transformed M = 1:37, SD = 0.97), $t(30)$ = 2.19, p = .037.

Lessons Learned

Because many of the results were comparable between the *technology as finder* and *as tool* approach, the lessons learned were likewise similar. However, there were a few notable differences that point to both the shortcomings and promise of the *tool* approach.

Positioning the technology as a tool narrowed its usefulness. A few visitors did enjoy using the tool: "*[I] like the magnifying glass (Beat Detector) … because I have to find it myself like a real scientist* (V27)." However, giving visitors the agency to search did little for those who had no idea of where or how to look for the heart at all. These were the very visitors who needed the most help. The tool's usefulness then narrowed to boosting the confidence of visitors who already knew to some degree, where the heart was.

Image recognition should clearly communicate and support its primary function. Given visitors' short time at any one exhibit and the sophisticated instrumentation they are learning to control, a tool with multiple applications may likely be used once for one targeted purpose. Furthermore, that key function needs to be perceived as relevant and essential to visitors' explorations. In our case, visitors largely saw the *Beat Detector* as a means for locating the heart. Any other purpose it could have served was secondary at best. Framing the image recognition as a tool with applicability beyond finding the heart diluted its usefulness for the visitors who simply wanted confirmation that they found the heart.

Framed as a tool, image recognition may help prolong exploration. Alternatively, there was some evidence that incorporating the technology as a tool helped sustain engagement; visitors who found the heart with the *Beat Detector* stayed longer, almost 30 seconds more, than those who did not. (We did not detect an equivalent difference in the *technology as finder* approach, possibly because it did not require visitors to interact to the same degree as the *technology as tool*.) This suggests that for a subset of visitors, the tool offered a means for prolonging their exploration.

CONCLUSION

The work described in this paper looked at three different approaches to integrating bioimaging techniques, specifically image recognition, into an interactive microscope exhibit with living specimens. Overall, this study found the following design lessons:

(1) Our use of image recognition was limited by the suboptimal quality of images (i.e. out of focus, no specimen in view) generated by visitors who were unfamiliar with what to look for and how to use the microscope. This precluded the technology's ability to definitively identify a specimen and control dynamic content as the specimen moved and changed, and to point out features for visitors who struggled to bring an image into focus. While better algorithms may be utilized, how well any image recognition technique will work is limited by visitors' ability to generate images for that technique to act on.

(2) Techniques outside of image recognition were needed. In our case, careful design and coordination between the biological characteristics of the specimen (i.e. predictable development trajectory) and materials (i.e. gas permeable gaskets) helped constrain the search space of what could and should be visible to increase the odds of quality input images for our algorithms.

(3) Even with a constrained search space, image recognition was not infallible, and the integration of the technology relied on framing it as an aid to exploration, specifically to confirm what visitors were looking at and, to a lesser extent, to point out something new. When presented as best, but imperfect, guesses to where to find interesting features (e.g., the heart), the technology helped bolster visitors' confidence in what they saw, which was valuable to visitors who were using authentic scientific instrumentation for the first time. For those visitors who struggled with finding interesting features, positioning the technology as a finder that would locate those features versus a tool, which they controlled to actively search for a feature, seemed to better support their explorations.

This case study focused on helping visitors investigate the microscopic world. We hope that its lessons may find applicability beyond microscopes to other science exhibits featuring dynamic phenomena, where image recognition may play a role in supporting visitors' explorations with a scientific lens.

ACKNOWLEDGMENTS

The authors wish to thank Rachel Cusing, Meghan Kroning, Tamara Kubacki, Rosario Soleto, and Jacqueline Wong for their evaluation assistance. We'd also like to thank Nenad Amodaj and Bing Dai for their technical advice. This material is based upon work supported by the National Science Foundation under grant 1514612. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. This work was also funded in part by the Gordon and Betty Moore Foundation.

REFERENCES

[1] Sue Allen. 2004. Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education* 88, S1: S17–S33. https://doi.org/10.1002/sce.20016

[2] Sue Allen and Joshua Gutwill. 2004. Designing with multiple interactives: Five common pitfalls. *Curator: The Museum Journal* 47, 2: 199–212. https://doi.org/10.1111/j.2151-6952.2004.tb00117.x

[3] Minda Borun and Jennifer Dritsas. 1997. Developing family-friendly exhibits. *Curator* 40, 3: 178–192.

[4] Josephine P. Briggs. 2002. The zebrafish: a new model organism for integrative physiology. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology* 282: R3–R9.

[5] Charles Carlson. 2005. Accessing the microscopic world. *PLoS Biology* 3, 1: e12. https://doi.org/10.1371/journal.pbio.0030012

[6] Karen R. Dell and Ronald D. Vale. 2004. A tribute to Shinya Inoué and innovation in light microscopy. *The Journal of Cell Biology* 165, 1: 21–26. https://doi.org/10.1083/jcb.200403023

[7] Gary Dodson and Cory Murphy. 2012. Zoo and aquarium webcams: An informed view. *Zoo Biology* 31, May 2011: 414–425. https://doi.org/10.1002/zoo.20405

[8] Richard O. Duda and Peter E. Hart. 1972. Use of the Hough Trasformation to detect lines and curves in pictures. *Communications of the ACM* 15, Jan 1972: 11–15. https://doi.org/10.1029/97JC01225

[9] Arthur Edelstein, Nenad Amodaj, Karl Hoover, Ron Vale, and Nico Stuurman. 2010. Computer control of microscopes using µmanager. *Current Protocols in Molecular Biology*, Suppl. 92: 1–17. https://doi.org/10.1002/0471142727.mb1420s92

[10] Kevin W. Eliceiri, Michael R. Berthold, Ilya G. Goldberg, Luis Ibáñez, B. S. Manjunath, Maryann E. Martone, Robert F. Murphy, Hanchuan Peng, Anne L. Plant, Badrinath Roysam, Nico Stuurman, Nico Stuurmann, Jason R. Swedlow, Pavel Tomancak, and Anne E. Carpenter. 2012. Biological imaging software tools. *Nature Methods* 9, 7: 697–710. https://doi.org/10.1038/nmeth.2084

[11] Kirsten Ellenbogen, Toni Dancu, and Cheryl Kessler. 2004. *Microscope Imaging Station Summative Evaluation*. Retrieved April 7, 2019 from https://www.exploratorium.edu/sites/default/files/pdfs/MIS_Summative.pdf

[12] John H. Falk. 2001. *Free-choice Science Education: How We Learn Science Outside of School*. Teachers College Press, New York.

[13] John H. Falk, Carol Scott, Lynn Dierking, Leonie Rennie, and Mika Cohen Jones. 2004. Interactives and visitor learning. *Curator: The Museum Journal* 47, 2: 171–198. https://doi.org/10.1111/j.2151-6952.2004.tb00116.x

[14] Christian Heath and Dirk Vom Lehn. 2008. Configuring “interactivity”: Enhancing engagement in science centres and museums. *Social Studies of Science* 38, 1: 65–93. https://doi.org/10.1177/0306312707084152

[15] George E. Hein. 2003. *Traits of Life*. San Francisco, CA. Retrieved April 7, 2019 from https://www.exploratorium.edu/sites/default/files/pdfs/Traits_summative.pdf

[16] Shinya Inoué. 1981. Video image processing greatly enhances contrast, quality, and speed in polarization-based microscopy. *The Journal of Cell Biology* 89, May: 346–356.

[17] Denise King, Joyce Ma, Angela Armendariz, and Kristina Yu. 2018. Developing interactive exhibits with scientists: Three example collaborations from the life sciences collection at the Exploratorium. *Integrative and Comparative Biology* 58, 1: 94–102. https://doi.org/10.1093/icb/icy010

[18] Seung Ah Lee, Engin Bumbacher, Alice M. Chung, Nate Cira, Byron Walker, Ji Young, Barry Starr, Paulo Blikstein, and Ingmar H. Riedel-kruse. 2015. Trap it ! : A playful human-biology interaction for a museum installation. In *Proceedings of the Conference on Human Factors* (CHI ’15), 2593–2602. https://doi.org/10.1145/2702123.2702220

[19] Jeff W. Lichtman and José-Angel Conchello. 2005. Fluorescence microscopy. *Nature Methods* 2, 12: 910–919. https://doi.org/10.1002/9783527687732.ch3

[20] Joyce Ma. 2016. *Seeing Scientifically Prototype Iteration 0 - Baseline Version*. San Francisco. Retrieved April 7, 2019 from https://www.exploratorium.edu/sites/default/files/pdfs/ssciBaseline_rp_02.pdf

[21] Joyce Ma and Jacqueline Wong. 2004. *Visitor Conversations around an Interactive Microscope Exhibit*. San Francisco, CA. Retrieved April 7, 2019 from https://www.exploratorium.edu/vre/pdf/conversation_rp_02.pdf

[22] Mahshid Majd and Reza Safabakhsh. 2017. Impact of machine learning on improvement of user experience in museums. *19th CSI International Symposium on Artificial Intelligence and Signal Processing, AISP 2017*: 195–200. https://doi.org/10.1109/AISP.2017.8324080

[23] National Research Council. 2009. *Learning Science in Informal Environments: People, Places and Pursuits*. National Academies Press, Washington D.C.

[24] National Research Council. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* National Academies Press, Washington D.C.

[25] Christiane Nüsslein-Volhard. 2012. The zebrafish issue of Development. *Development* 139, 22: 4099–4103.

[26] Jessie Pallud. 2017. Impact of interactive technologies on stimulating learning experiences in a museum. *Information and Management* 54, 4: 465–478. https://doi.org/10.1016/j.im.2016.10.004

[27] Randi Korn & Associates. 2003. *The Science Museum of Minnesota Cell Lab Summative Evaluation*. Retrieved April 7, 2019 from https://www.smm.org/sites/default/files/public/attachments/cell-lab.pdf

[28] Cody Sandifer. 2003. Technological novelty and open-endedness: Two characteristics of interactive exhibits that contribute to the holding of visitor attention in a

science museum. *Journal of Research in Science Teaching* 40, 2: 121–137. https://doi.org/10.1002/tea.10068

[29] Leona Schauble and Karol Bartlett. 1997. Constructing a science gallery for children and families: The role of research in an innovative design process. *Science Education* 81, 6: 781–793.

[30] Caroline A. Schneider, Wayne S. Rasband, and Kevin W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9, 7: 671–675. https://doi.org/10.1038/nmeth.2089

[31] Beverly Serrell and Roxana Adams. 1998. *Paying Attention: Visitors and Museum Exhibitions*. American Association of Museums.

[32] Lisa F. Smith, Jeffrey K. Smith, Kimberly K. Arcand, Randall K. Smith, and Jay A. Bookbinder. 2015. Aesthetics and astronomy: How museum labels affect the understanding and appreciation of deep-space images. *Curator: The Museum Journal* 58, 3: 283–297.

[33] Leonard Wein. 2014. Visual recognition in museum guide apps. In *Proceedings of the Conference on Human Factors* (CHI ’14), 635–638. https://doi.org/10.1145/2556288.2557270