



METHOD ARTICLE

Sonification of hyperspectral fluorescence microscopy datasets [version 1; referees: 1 approved, 1 approved with reservations]

Aprameya Mysore¹, Andreas Velten^{1,2}, Kevin W. Eliceiri^{1,2}

¹Laboratory for Optical and Computational Instrumentation, University of Wisconsin, Madison, USA

²Morgridge Institute for Research, Madison, USA

v1 First published: 25 Oct 2016, 5:2572 (doi: [10.12688/f1000research.9233.1](https://doi.org/10.12688/f1000research.9233.1))
 Latest published: 25 Oct 2016, 5:2572 (doi: [10.12688/f1000research.9233.1](https://doi.org/10.12688/f1000research.9233.1))

Abstract

Recent advances in fluorescence microscopy have yielded an abundance of high-dimensional spectrally rich datasets that cannot always be adequately explored through conventional three-color visualization methods. While computational image processing techniques allow researchers to derive spectral characteristics of their datasets that cannot be visualized directly, there are still limitations in how to best visually display these resulting rich spectral data. Data sonification has the potential to provide a novel way for researchers to intuitively perceive these characteristics auditorily through direct interaction with the raw multi-channel data. The human ear is well tuned to detect subtle differences in sound that could represent discrete changes in fluorescence spectra. We present a proof of concept implementation of a functional data sonification workflow for analysis of fluorescence microscopy data as an FIJI ImageJ plugin and evaluate its utility with various hyperspectral microscopy datasets. Additionally, we provide a framework for prototyping and testing new sonification methods and a mathematical model to point out scenarios where vision-based spectral analysis fails and sonification-based approaches would not. With this first reported practical application of sonification to biological fluorescence microscopy and supporting computational tools for further exploration, we discuss the current advantages and disadvantages of sonification over conventional spectral visualization approaches. We also discuss where further efforts in spectral sonification need to go to maximize its practical biological applications.

Open Peer Review

Referee Status:

	Invited Referees	
	1	2
version 1 published 25 Oct 2016	 report	 report
1	Paul S. Maddox , University of North Carolina at Chapel Hill USA	
2	Andrew R. Cohen , Drexel University USA	
Discuss this article		
Comments (0)		

Corresponding author: Kevin W. Eliceiri (eliceiri@wisc.edu)

How to cite this article: Mysore A, Velten A and Eliceiri KW. **Sonification of hyperspectral fluorescence microscopy datasets [version 1; referees: 1 approved, 1 approved with reservations]** *F1000Research* 2016, 5:2572 (doi: [10.12688/f1000research.9233.1](https://doi.org/10.12688/f1000research.9233.1))

Copyright: © 2016 Mysore A *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Competing interests: No competing interests were disclosed.

First published: 25 Oct 2016, 5:2572 (doi: [10.12688/f1000research.9233.1](https://doi.org/10.12688/f1000research.9233.1))

Introduction

Use of spectral information in microscopy

The increased availability of microscopes with multiple spectral channels and multi-colored fluorescent molecular markers has allowed life science researchers to generate datasets of higher spectral complexity than ever before from their optical imaging systems¹. By selectively placing different markers throughout a specimen, a researcher can construct detailed visual narratives of different biological processes based on the spectral variations throughout the data²⁻⁵. This approach, called multiplexing, allows for functional readout of multiple genetic functions spatially and temporally, a feature unique to fluorescence imaging. These multiple fluorophores can be genetically engineered extrinsic fluorescence tags such as Green Fluorescent Protein (GFP) or endogenous fluorescence such as that emanating from cellular autofluorescence. In particular, spectral information has been used to discriminate fluorescence signals from cellular autofluorescence signals such as nicotinamide adenine dinucleotide (NADH) in metabolism studies and to understand subcellular protein trafficking via signaling protein labeling^{6,7}. Additionally the collection of spectral data enables researchers to study other complex phenomena related to fluorescence such as fluorescence resonance energy transfer (FRET)⁸ or spectral shifts.

Acquisition and interpretation of spectral data

As fluorescence microscopy datasets increase in richness and dimensional complexity, designing flexible tools for researchers to explore these larger and denser datasets that strike the appropriate balance between intuitive functionality and analytic effectiveness is becoming increasingly important.

As the number of fluorescent markers used in a biological sample increases, two critical bottlenecks emerge within the current traditional analysis workflow. First, the spectral properties of fluorescent dyes often necessitate a linear unmixing (deconvolution) of the image data to discriminate one dye from the other. The importance of this step becomes particularly apparent when simultaneously using many fluorophores with overlapping emission spectra. Several types of spectral stripping and unmixing techniques have been borrowed from other scientific domains such as space science that involve analysis of hyperspectral data^{9,10}. However, many of these techniques require precise a priori knowledge of the present spectra before the imaging^{5,11-13}. Confirming emission spectra in experiments involving biological specimens becomes complicated due to factors such as spectral variance in endogenous chemical environment, the developmental fluorophore stage, and unpredictable refractive characteristics inside the specimens. All can cause significant variation in emission spectra¹⁴.

Second, even if all spectra are successfully unmixed, visual representations of an image volume with more than three spectral channels are limited by inability of the human vision system to manage effectively more than three channels at a time. False color images of datasets with three or more spectral dimensions can visually appear to be the same in areas where the data are vastly different, because the transcoding operation used to render the raw data visible to human eyes involves a data loss. Therefore visual analysis of hyperspectral imagery has to always rely on analysis of multiple images or on custom computational pre-processing.

While the human vision system is limited for natively comprehending rich hyperspectral data, the human ear and auditory system are more optimally suited for this task¹⁵. **Figure 1** shows a spectral sensitivity curve for the human eye, while **Figure 2** shows the “equal loudness curve” describing the range of frequencies perceptible by the human ear. Note the logarithmic scaling on the x-axis of **Figure 1** as compared to the relatively narrow interval on the x-axis of **Figure 2**. Additionally, the eye has only three channels of spectral sensitivity, the ear can distinguish about 1400 pitches throughout its range¹⁶. Moreover, the ear actually has finer resolution when sensing complex, non-sinusoidal tones¹⁵. This behavior makes the ear a substantially more optimal sensor for spectrally rich signals than the eye, because, as we demonstrate,

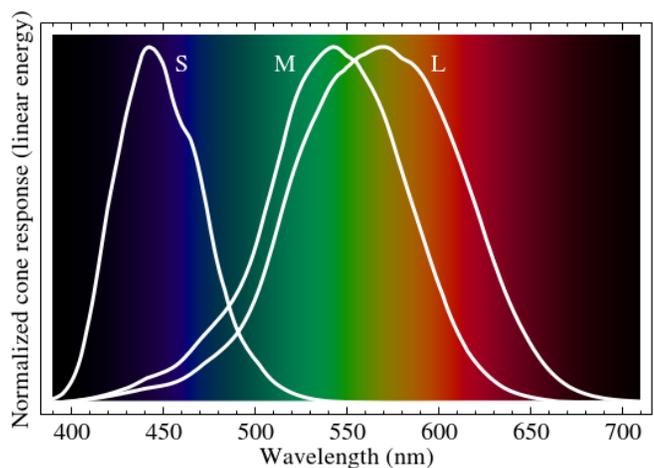


Figure 1. Spectral sensitivity of human eye. This chart shows the combined spectral sensitivity of all the color cones in the human eye, which constitutes the full range of visually perceptible frequencies.

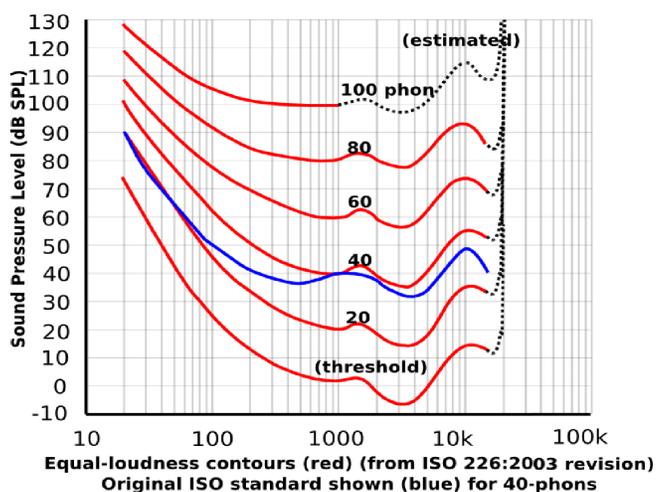


Figure 2. Spectral sensitivity of human ear. This chart shows the spectral sensitivity of the human ear in the form of multiple equal loudness contours over a frequency domain. While the eye has a domain of several hundred Hz, the ear has domain spanning several thousand Hz, as seen by the log scaling.

in [Figure 1](#) and [Figure 2](#), the eye actually becomes less spectrally adept with increasing spectral complexity. We intend to show how sonification allows researchers to avoid complex computational and visual techniques to best interpret their hyperspectral datasets.

Sonification as exploratory data analysis

Research on data sonification has been conducted across a diverse variety of disciplines and has yielded results across the spectrum from purely aesthetic to highly functional. The emphasis in our project was on creating a sonification plugin as a functional and practical tool for fluorescence microscopy research. As such, we decided to design our sonification as an extension of the imaging and analysis workflow that researchers at our lab use to gather and understand multispectral fluorescence microscopy data. Sonification has been proven a viable means for analyzing hyperspectral or otherwise dimensionally complex datasets in many other contexts¹⁷⁻¹⁹. Potential microscopy sonification applications have been briefly explored but not actively pursued and implemented by the scientific community^{20,21}. Dr. Thomas Hermann, in his 2002 PhD thesis “Sonification and Exploratory Data Analysis”, considers the viability of a hypothetical sonification-based multichannel image analysis tool for microscopy research²². In his thought experiment he outlines three ways that sonification would uniquely provide additional clarity and insight into dimensionally complex microscopy datasets:

--Sonification intuitively allows users to perceive trend-related factors such as clustering, or other distributional behavior with increased clarity by enunciating more subtle spectral differences that missed by more conventional data display methods. This allows users to intuitively understand higher order relationships amongst the 15+ channels that might go unnoticed when visually exploring the image stack 1–3 channels at a time.

--Sonification augments perception when the eyes are saturated with other information. If a user needs a visual interface to explore the structural organization of a particularly intricate specimen, the sonification can simultaneously provide information on salient characteristics from other channels co-localized to a particular region of interest without obfuscating the visual-spatial information.

--Sonification enhances the salience of ‘warning flags’. Rapid changes in sound naturally command the human attention very effectively. It is a lot easier to grasp the difference in amplitude between two points in the dataset than it is to compare the intensity of two points defined by multiple overlaid color maps.

Currently, no data sonification package exists for hyperspectral image volume exploration and analysis that avoids computational dimensionality reduction. Researchers at Stanford have implemented a sonification model for hyperspectral image analysis, but this relies heavily on principal component analysis and avoids sonification of the raw multichannel pixel stream²¹. Our approach uses data sonification to reduce the amount of abstraction between the researcher and the raw fluorescence data. We explore sonification as a way of confirming a spectral deconvolution, as well as a way of navigating the fluorescence data in its full dimensionality. By allowing researchers to interact directly with

the spectral composition of their datasets, we can potentially elucidate salient characteristics of the data that otherwise would have gone unnoticed.

Materials and methods

There are two components to the sonification. The first is the image-viewing interface that provides the user interface for the data to be explored and sonified. For this we chose a plugin context inside FIJI (<http://fiji.sc/>) which is a popular distribution of the Java-based image analysis program ImageJ^{23,24}. Choosing an open-source image analysis package was important to us as to allow maximum accessibility and ease adoption. FIJI is already the tool of choice for many scientists doing microscopy research. Several of the data sonification implementations described in the current research literature use a standalone application which integrates image UI and audio driver communication into one executable, but this was not the optimal strategy for our proposed use case. A FIJI plugin was the intuitive choice for development and deployment because it can easily be integrated into current research workflows instead of requiring the installation and usage of a separate standalone application. Additionally, since FIJI is used for many spectral fluorescence microscopy analysis routines, it opens up the possibility of integrating and interacting with these steps as needed by the end user.

The second component is the audio server, which is tasked with storing synths and outputting sounds to the native audio driver. Synths are definitions that describe the range of sounds produced by the sonification, and contain instructions on how to translate the commands received from the image host into audio samples. As we discuss in the “Sound Design” section, iteratively designing multiple mappings from data-space to sound space is how we explore the capacity of sonification to allow us to interact with hyperspectral datasets. We decided to use SuperCollider (<http://supercollider.sourceforge.net/>) and JCollider (<http://www.sciss.de/jcollider/>) for the functionality of storing synths and communicating with the audio driver. SuperCollider is an open-source package for algorithmic sound synthesis, and JCollider is a library that allows Java applications, such as ImageJ, to communicate with an instance of SuperCollider’s OpenSoundControl (OSC) server, called “scsynth”. This communication entails encoding and sending audio commands, manipulating synth definitions, and checking server status. Synth definitions are encapsulated in SynthDef Objects, as specified in the JCollider API, and exported to scsynth at runtime. All communications between the Java image host and the OSC server are done over User Datagram Protocol (UDP).

The following signal flow diagram in [Figure 3](#) describes the interaction process of the plugin. The user interacts with the image via the FIJI interface by clicking and dragging the cursor across the image, which sends pixel data encoded in OSC format to scsynth. The data are then sonified by the SuperCollider server in accordance with the chosen synth definition and sent to the speakers as audio samples.

Sound design methodology

When designing the mapping from data space to sound, we first had to establish a set of heuristics to determine what qualities of the sound are desirable in a particular mapping. Grond and Berger’s

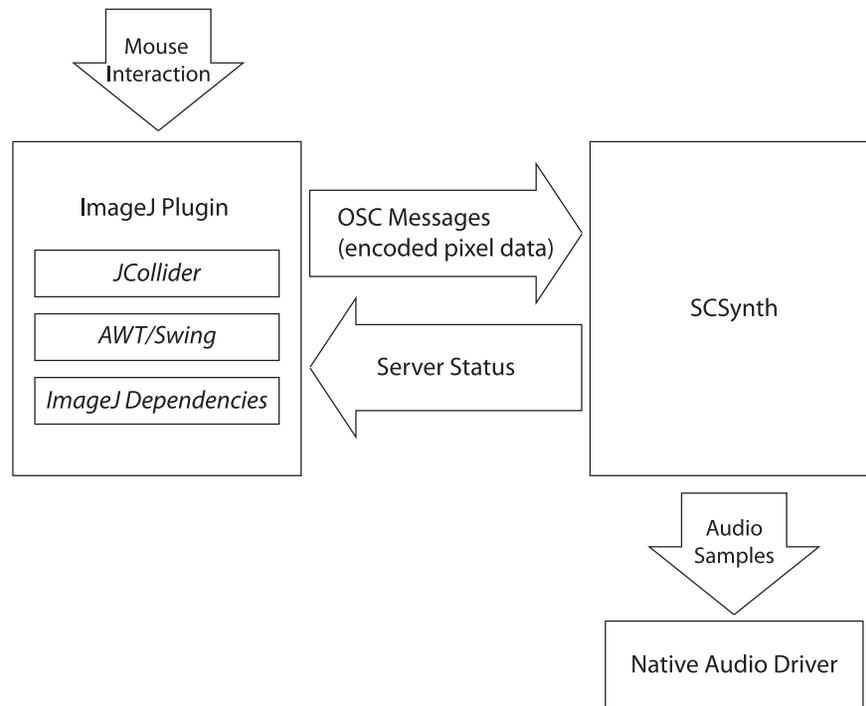


Figure 3. Block diagram of sonication plugin. This signal flow diagram describes the data sonification pipeline. When the user interacts with the image volume by clicking and dragging, FIJ sends encoded pixel values to the SuperCollider server, which translates the data into audio samples sent to the speakers.

chapter in the “Sonification Handbook”, entitled “Parameter Mapping Sonification”, informed our sound design methodology²⁵. They outline the following optimal design criteria for a parameter based exploratory sonification:

Polarity. Polarity constitutes the capacity of a sound to communicate both the direction and magnitude of a change in a stream of data. In the use-case of sonifying a spectrally rich signal, as an individual component of the set of spectra increases or decreases, such as the pixel intensity value for a particular channel in an image volume, the respective mapped quantity of the sound should change in a way that can be clearly identified as the increase or decrease.

Scale. Data from our image volumes have to be appropriately scaled in the mapping process in order to convey the full range of parameter fluctuations. Ideally the same image in two different formats should sound the same, so normalization of the pixel stream data should be implemented to minimize arbitrary differences in auditory characteristics due to image encoding and prevent signal spikes. Images encoded in a format with higher dynamic range have higher amounts of noise though, so there will invariably be some differences.

Context. Context is the auditory equivalent of “axes and tick marks”—elements of the sonification that provide a baseline reference against fluctuations in the data²⁵. If fluctuations appear to be arbitrary or ambiguously connected to observed changes in

the data, the utility of the sonification as an exploratory mechanism is severely diminished. As such, we sought to design our synths with as little auditory clutter as possible; one should be able to selectively focus on particular modulating qualities in the sound and intuitively know how the data are changing. This rule of thumb definitely informed our preference for additive synthesis. Another issue related to context we observed was that, similar to the case with color lookup tables, there is no “one size fits all” auditory display solution. Certain synth designs are more useful for articulating subtle changes in spectral composition but ineffective for perceiving larger trends across the entire dataset, and *vice versa*. As such, we provide a variety of synths, each for usage in a specific observed context.

Calibration. We employ a series of calibration images designed to systematically screen any arbitrary synthesis algorithm for viability and identify what spectral contexts are or are not displayed clearly by a given sound. Figure 4 shows a comparison of the four volumes we generated. Each calibration image volume consists of 15 450px by 450px TIFF images representing the spectral data shown in Figure 5a–Figure 5d.

Note in Figure 4 there is a 30 pixel thick horizontal stripe that traverses the entire width of the image in each of the volumes. This stripe starts at the top of the image in channel 1 of each volume, and incrementally shifts downward 30 pixels each channel until it is at the very bottom of the image in channel 15, as seen in Figure 5a–Figure 5d. Each of these stripes has gradation from 8-bit

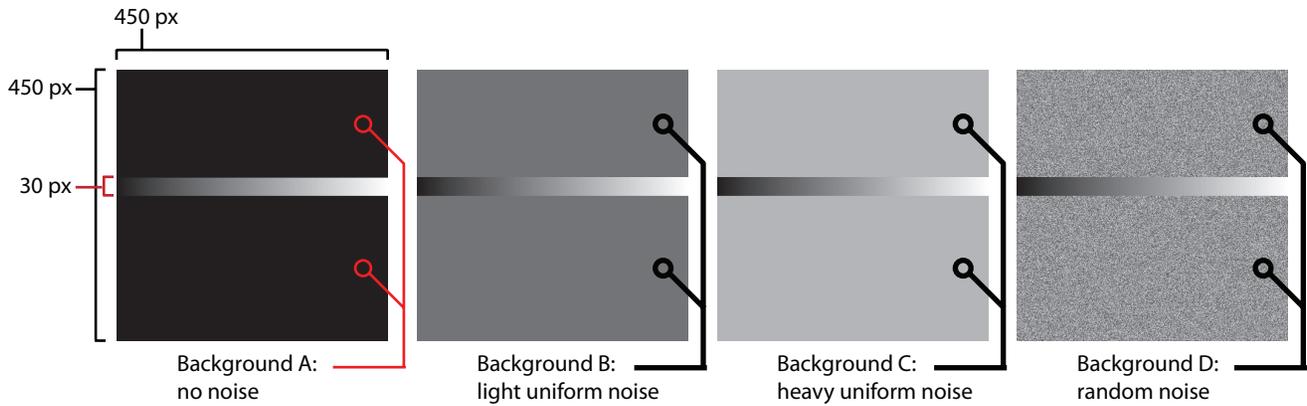


Figure 4. Comparison of calibration image volumes. Single channel slices of each of the four calibration image volumes are compared side by side. In each volume, the same cascading gradient is interpolated with a different 'background noise' texture. This particular design of a calibration image allows for testing the signal-noise discrimination capabilities of a given synth for each channel in a variety of spectral contexts.

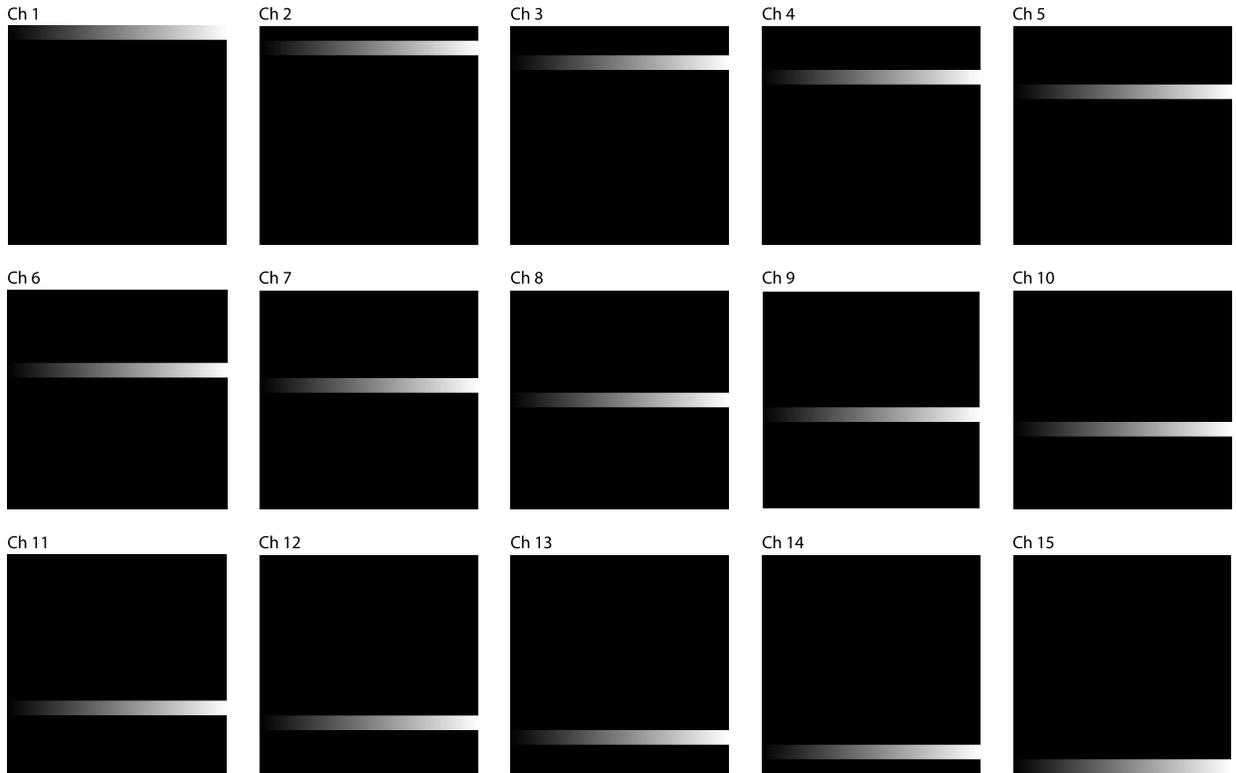


Figure 5a. Calibration Volume "A"--No background noise. Calibration Volume 'A' contains the cascading 30px gradient with no background noise. This is a control volume to test a particular synth's range of sounds with no background interference.

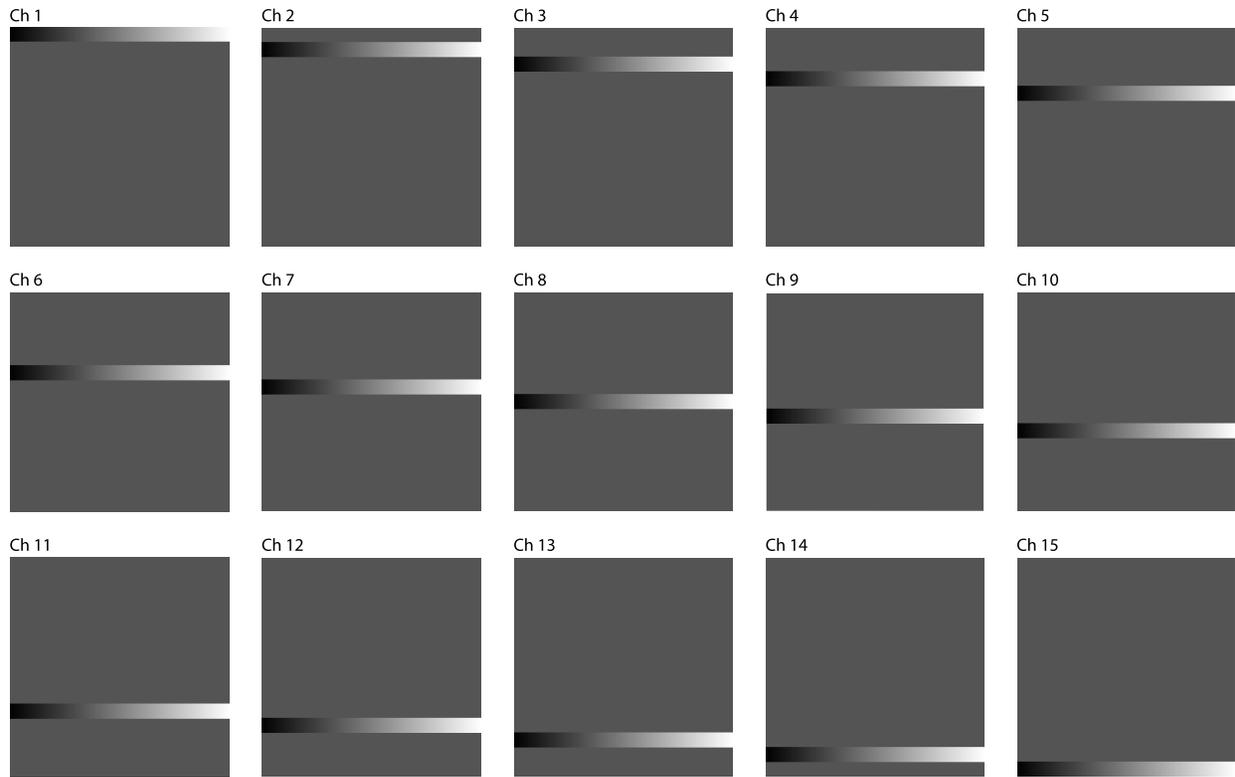


Figure 5b. Calibration Volume “B”--Light uniform background noise. Calibration Volume 'B' contains the cascading 30px gradient with low background noise. This volume tests a particular synth's signal-noise discrimination with low intensity homogeneous background noise across all channels.

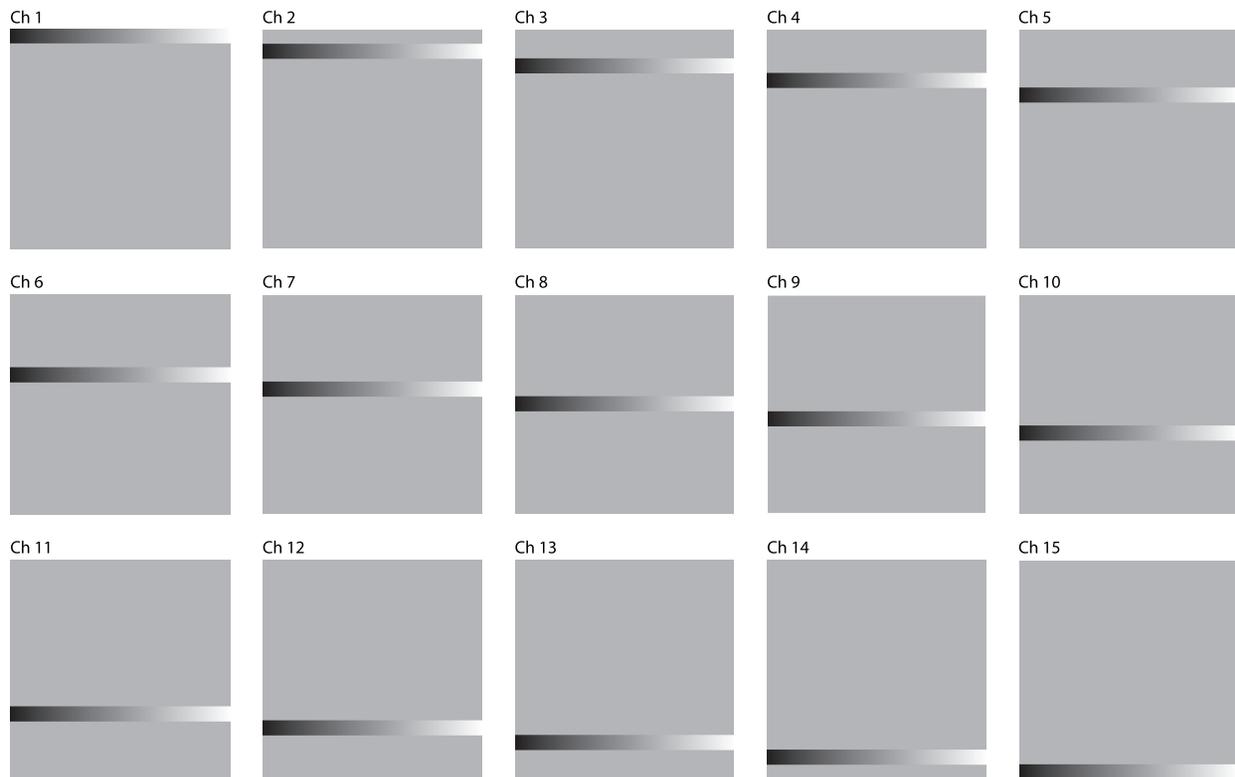


Figure 5c. Calibration Volume “C”--Heavy uniform background noise. Calibration Volume 'C' contains the cascading 30px gradient with high background noise. This volume tests a particular synth's signal-noise discrimination with high intensity homogeneous background noise across all channels.

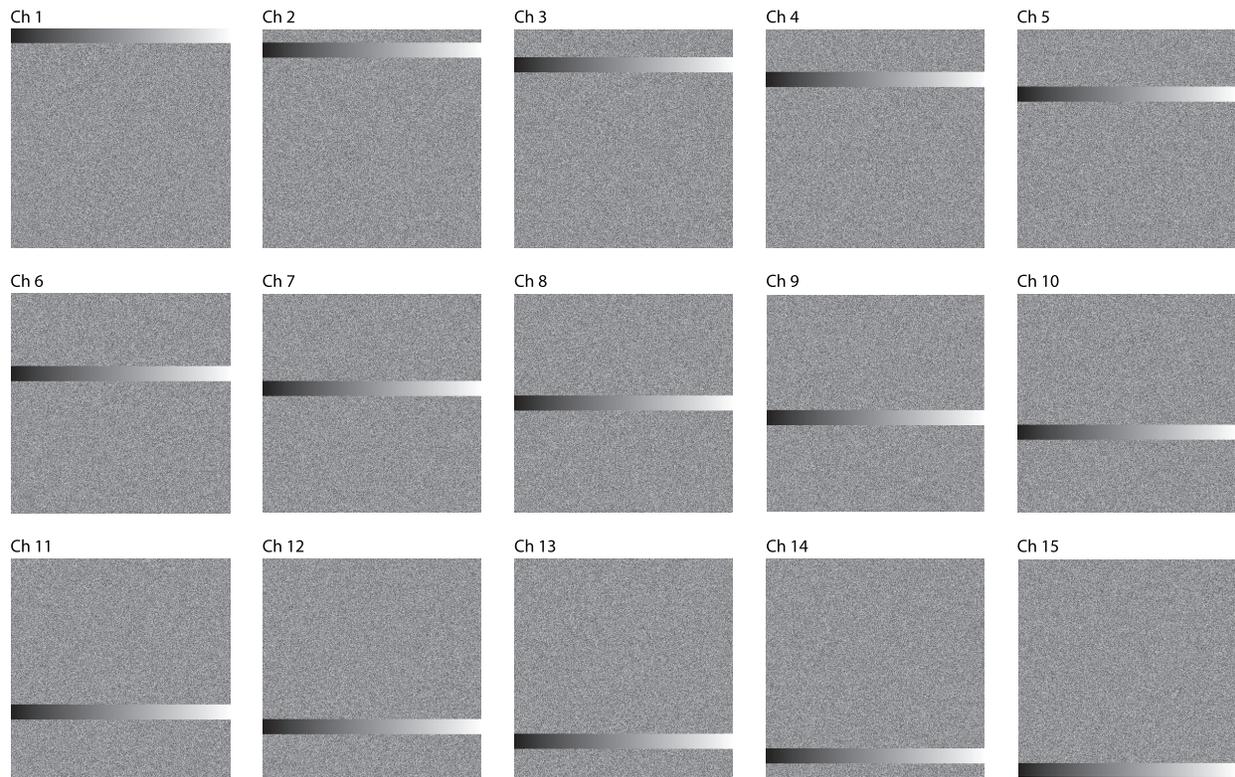


Figure 5d. Calibration Volume “D”--Random background noise. Calibration Volume 'D' contains the cascading 30px gradient with random background noise. This volume tests a particular synth's signal-noise discrimination with pseudo-random, heterogeneous background noise across all channels.

grayscale intensity value 0 (black) on the leftmost side to 255 (white) on the rightmost side. Throughout the rest of the image not occupied by the stripe, each of the four calibration volumes has a different pattern, designed to simulate various types of noise that could interfere with the perception of the 0–255 gradient. Volume A is designed to be the control volume which should showcase the discriminatory function of a sonic mapping at its clearest. Volumes B and C both have noise that is homogenous throughout the entire background; B has light noise (grayscale value = 84/33% intensity) while C has heavier noise (grayscale value = 168/66% intensity). Volume D has pseudorandom noise that is clamped between 20% and 80% intensity (to prevent unnecessary clipping).

When the 15 images of a calibration volume are superimposed on top of one another in an image stack, the cascading stripe and background pattern provides a means to measure the effectiveness of a particular sonic mapping at discriminating intensities for any given channel by traversing the image horizontally. Conversely, the spectral differentiation of the mapping can be gauged at a given intensity by traversing the image vertically. [Movie 1](#) shows how sonifying these four volumes with a given sound can give an idea of its effectiveness in a variety of spectral contexts.

Video of Calibration

1 Data File

<http://dx.doi.org/10.6084/m9.figshare.4029480.v1>

Synth definitions

Each synth definition we created is outlined here as a list of summations defining how audio samples. These following synths were utilized on 15 channel datasets, so there are 15 components to each synth waveform that are modulated by fluctuations in the spectral content of a given stream of pixels. Additive synthesis is the primary algorithm used to group all these fluctuations into one sound in the definitions we created.

Synth 1.

$$y_1(t) = \sum_{a=1}^{15} i_a \sin(2\pi f_a t)$$

$f = \{440\text{Hz}, 466.16\text{Hz}, 493.88\text{Hz}, 523.25\text{Hz}, 554.37\text{Hz}, 587.33\text{Hz}, 622.25\text{Hz}, 659.26\text{Hz}, 698.46\text{Hz}, 739.99\text{Hz}, 783.99\text{Hz}, 830.61\text{Hz}, 880\text{Hz}, 932.33\text{Hz}, 987.77\text{Hz}\}$

i_a = normalized pixel intensity value (0–1) for channel ‘a’ of a multichannel image volume

This synth maps the intensities of spectral channels 1–15 to the amplitudes of sine waves from A4 to B5, and then adds all the waves together to produce the resultant sonification for a given pixel. We found that this synth design was effective for the phantom, but was less useful for articulating subtle spectral distinctions in biological datasets.

Synth 2.

$$y_2(t) = \sum_{a=1}^{15} i_a \sin(2\pi 300at)$$

This synth uses the same additive principle of Synth 1 (intensities mapped directly to amplitude), but for a different set of notes. Instead of incremented tones on a scale, this synth uses a harmonic series. We found this synth to be notably more useful when sonifying biological datasets than Synth 1 as the wider range of notes makes it more difficult for the user to mistake adjacent spectral channels for one another.

Synth 3.

$$y_3(t) = \sum_{a=1}^{15} 0.1 \sin(2\pi(300a + 150i_a)t)$$

This mapping is also an additive synthesis algorithm, but instead of modulating the amplitudes of the components, the actual frequencies are modulated to create inharmonics. The amplitude of each component is set to 0.1 (statically), and the intensities of spectral channels 1–15 are mapped to a range of 0–150Hz deviation from each harmonic partial.

Synth 4.

$$y_4(t) = \sum_{a=1}^{15} i_a \sin(2\pi(300a + 150i_a)t)$$

This mapping combines modulation of amplitude with deviation from harmonic partials. We found the combination of the two allowed for better spectral differentiation than either Synth 1 or Synth 3 alone.

Synth 5.

$$y_5(t) = \sum_{a=1}^{15} i_a \sin(2\pi((300 + 600(a-1)) + 350i_a)t)$$

This is the mapping we used in all the example videos. The summation is constructed the same way as Synth #4, except the range of notes is expanded by choosing every other higher order

harmonic partial of 300Hz. Additionally the frequency deviation from the harmonics, mapped to the pixel intensity, is scaled to 350Hz as opposed to just 150Hz. We found it to display subtle spectral changes with the most clarity out of all the synths.

Results

Dataset 1. Raw data for ‘Sonification of hyperspectral fluorescence microscopy datasets’, Mysore et al. 2016

<http://dx.doi.org/10.5256/f1000research.9233.d138682>

See README.txt for a description of the files.

Generation of hyperspectral phantom dataset

The human eye makes use of photoreceptors with overlapping spectral sensitivity in order to improve its spectral resolution. It is important to note that this method does not increase the total number of colors or color combinations the eye can perceive, but rather optimizes color perception in certain situations that are more likely to be relevant in a natural environment while decreasing resolution in other situations. Specifically, overlapping color receptors allow the eye to establish the mean frequency of a single emission peak with greater accuracy. When, however, multiple peaks are involved this method becomes unreliable since there are always different peak configurations that will lead to the same perceived color. To illustrate this effect below we present a simple method to generate those ambiguous peak configurations for any given photoreceptor sensitivities. We demonstrate this method by generating a dataset made up of five emitters. We produce a hyperspectral image in which every pixel has a different spectrum due to different relative fluorophore concentration or emission wavelengths, but all appear identical to a three channel detector with overlapping spectra such as the eye.

Consider a sensor with three overlapping channels detecting a spectrum consisting of five relevant channels. The input spectrum can be described as a five-element vector I and the detected intensities at the sensor are described by a three-element vector D . D can be computed from I by multiplying a 3 by 5 matrix M that in its columns contains the spectral sensitivity curves of the three spectral channels.

We consider D and M fixed and are seeking all vectors I that fulfill the equation

$$D = MI$$

By varying u and v we can create a set of five channel spectra that all result in the same vector D . This D can contain negative numbers, which is not realistic. To avoid this we add set of constants $\Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5$ to our spectra I such that the smallest value occurring is zero. This leads to a different detected spectrum D' that is still constant for all incident fluorophore spectra:

The result of this procedure, our 5-channel hyperspectral phantom dataset, is shown in [Figure 6](#). Each pixel in the image multichannel volume has a different 5-fluorophore combination. They all look identical to the three channels sensor used. The spectra do however sound dramatically different in a simple sonification.

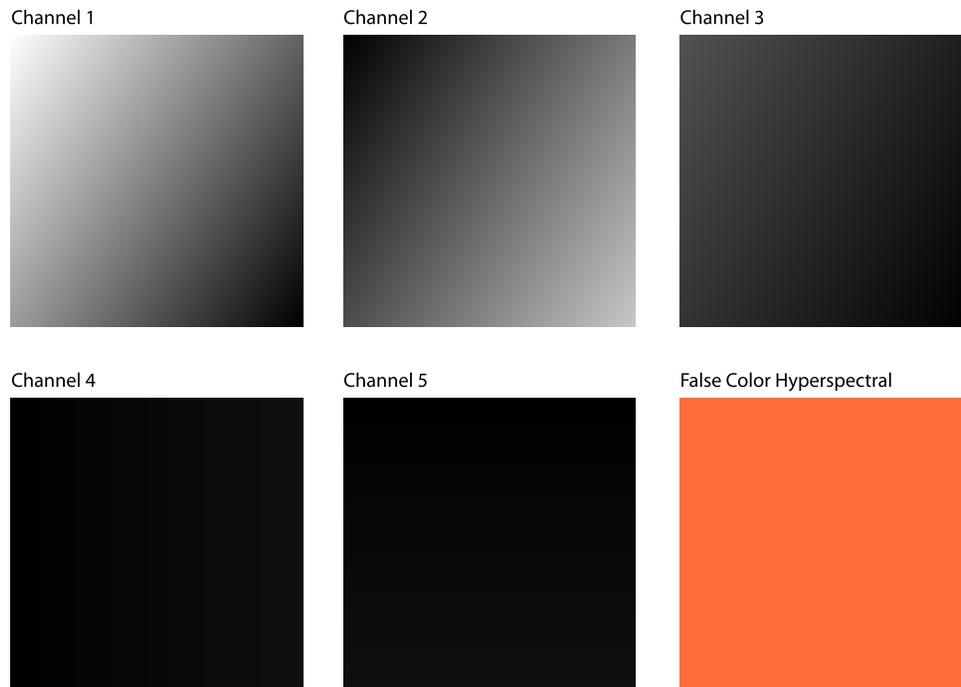


Figure 6. Phantom image volume. This figure shows a visual comparison of each channel of the phantom image volume alongside a hyperspectral false color image of the volume. While each channel has a unique gradient across the image, the hyperspectral false color image appears homogeneously textured throughout.

Sonification of hyperspectral phantom dataset

The sixth image seen in [Figure 6](#) is a hyperspectral false color image of our phantom image volume. This false color image is a mapping of the 5D data to 3D space, using the spectral sensitivity curve seen in [Figure 7](#), which has a spectral overlap similar to the cone sensitivity of the human eye shown earlier in [Figure 1](#).

Although the false color image appears to be uniform across its entire surface, the actual spectral data, shown on the left, are different for every pixel. Even though there is significant overlap between the spectral peaks used in the false color mapping, there are no apparent visual distinctions in the RGB color image. [Figure 8](#) shows the invisible variations in spectral content that characterize our image volume.

This data loss does not reflect an inadequacy in the sensitivity curve, but rather the fundamental inability of a three-dimensional transcoding to convey fluctuations in a five-dimensional signal. In the above section detailing the generation of the phantom dataset we describe how any arbitrary spectral response curve with three peaks can be used to compute a five-dimensional image volume that varies spectrally at each pixel, and a RGB false color hyperspectral image that appears to be spectrally homogeneous across all pixels to the human eye.

Using the sonification plugin, the user traverses the image volume and generates an audio signal that elucidates the invisible spectral topology of this image volume. A log on the left displays spectral data for each pixel that is sonified, confirming the spectral difference. This process is shown in [Movie 2](#).

Sonification of phantom image volume

1 Data File

<http://dx.doi.org/10.6084/m9.figshare.4029495.v1>

Sonification of FocalCheck beads

After demonstrating theoretical utility with the phantom dataset, we seek to create a test for our sounds that would more closely resemble the spectral subtleties of a real biological specimen with multiple fluorophores. However, we want a dataset that is much more spectrally controlled than a biological specimen so we can unequivocally verify our plugin's functionality.

The FocalCheck DoubleGreen Fluorescent Microspheres Kit (Life Technologies, Carlsbad CA) provides a useful litmus test for using the sonification to differentiate two objects with very similar spectra. The three different fluorescent beads in this kit are uniformly labeled and are designed to test the spectral separation capabilities of a scope system. The subtle distinction in dye coverage and color between the three beads is designed to be visually unnoticeable, but precisely pronounced enough to be resolved by algorithmic linear unmixing. By testing our sonification with the FocalCheck beads, we can at once ascertain the capability of our plugin at confirming the results of a linear unmixing as well as differentiating spectra that are visually indistinguishable.

[Figure 9](#) shows magnified false color images of the three beads with spectral measurements taken both at the core and the edge for each bead. The bead in [Figure 9a](#) homogeneously emits "green 1" (512 nm) and the bead in [Figure 9b](#) homogeneously emits "green

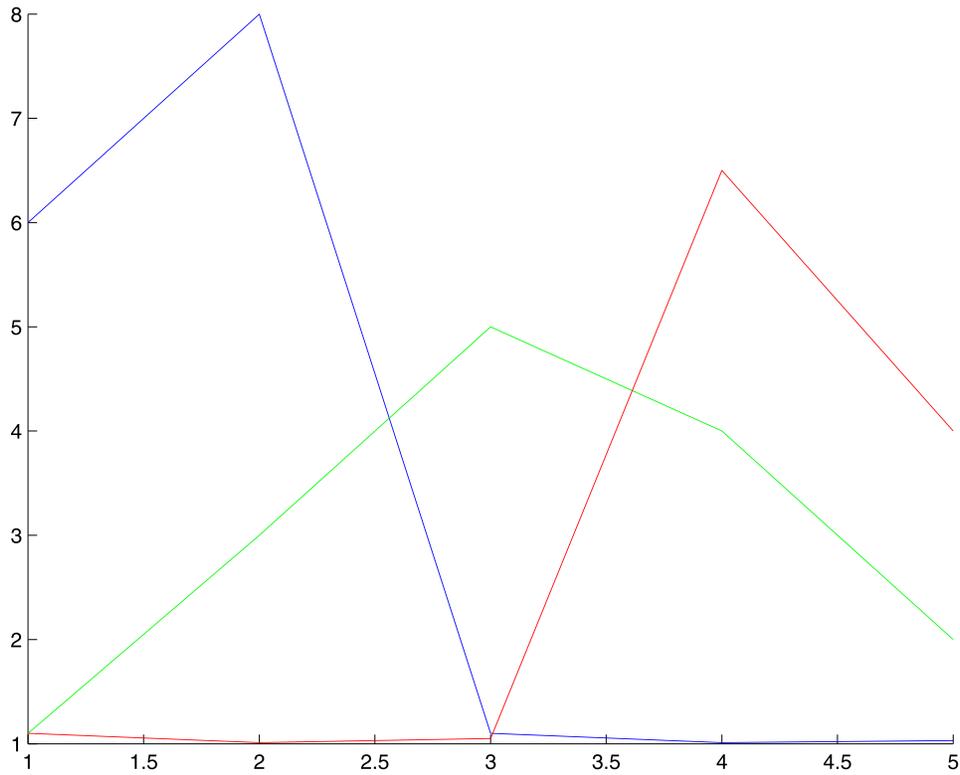


Figure 7. Assumed RGB Sensitivity Spectra. This figure shows the spectral sensitivity curve used to generate the hyperspectral false color image from the phantom image volume. In this curve the three color components are overlapping to emulate the sensitivity curve shown in Figure 1.

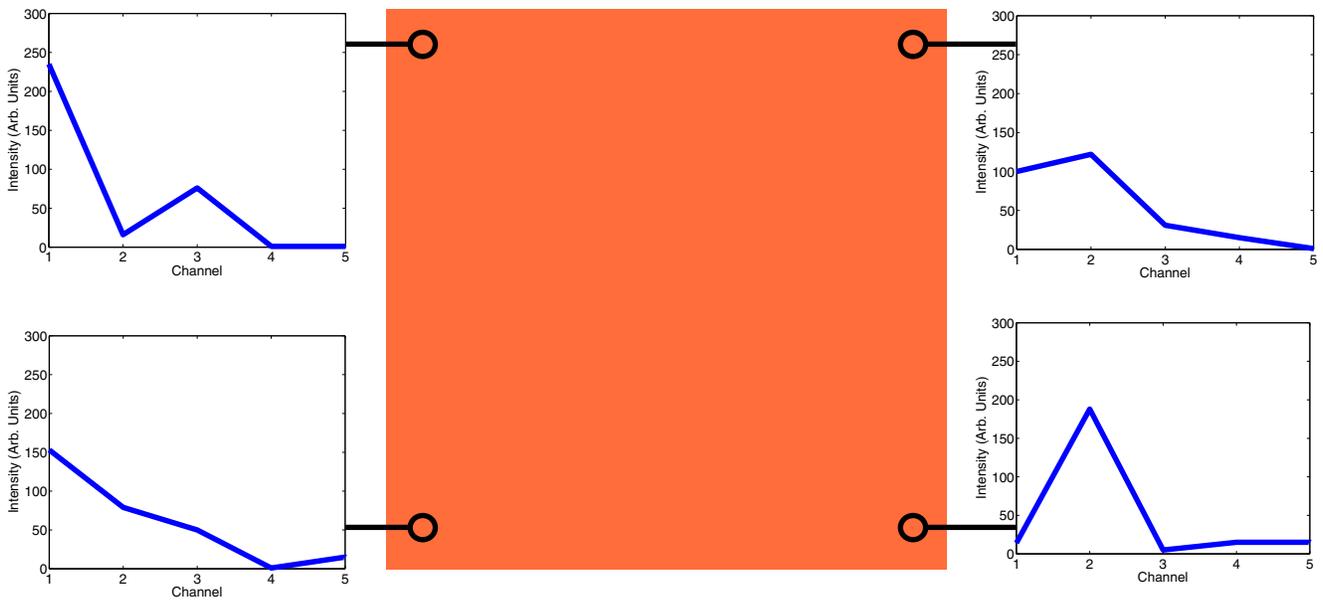


Figure 8. Spectral characteristics of phantom image volume. The spectral composition of the phantom image volume is shown at various points throughout the volume and overlaid with the hyperspectral false color image. This shows the range of spectral variations that are not visible through the false color image.

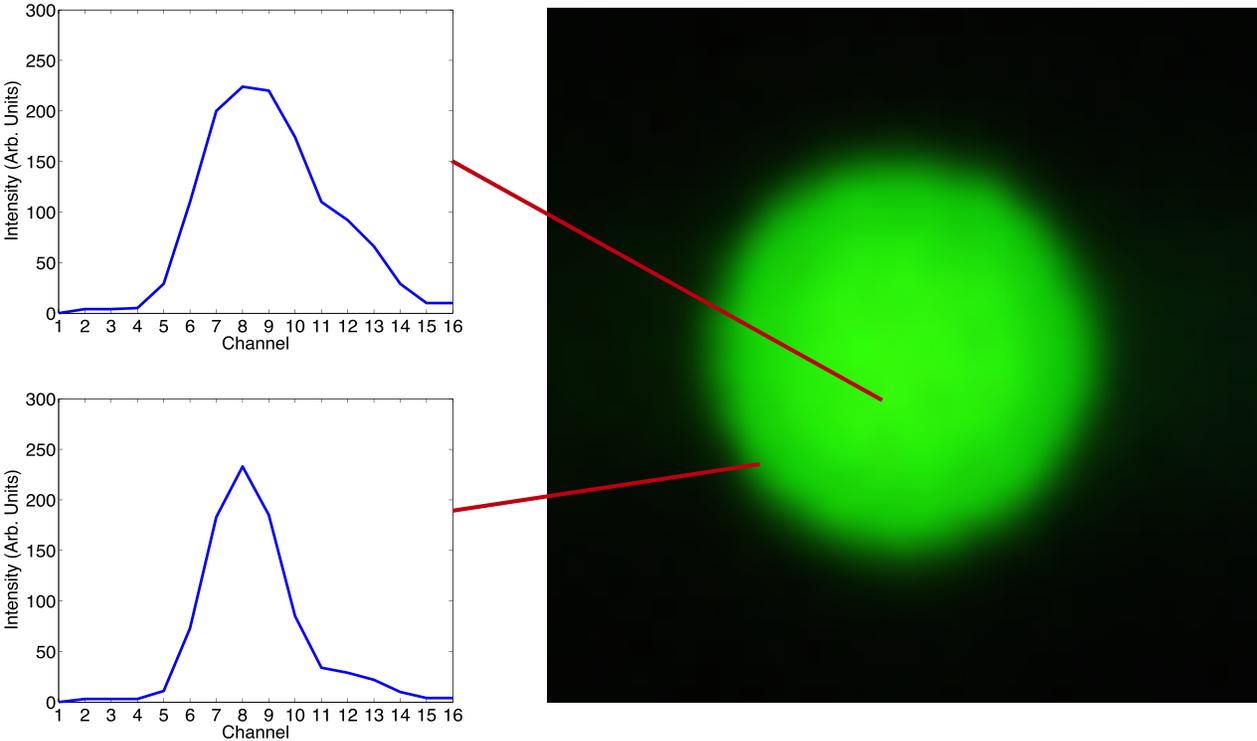


Figure 9a. Comparison of core vs. edge of FocalCheck bead: Green 1. The spectral content of the FocalCheck DoubleGreen bead 1 (512nm emission) is shown at the core and at the edge of the bead.

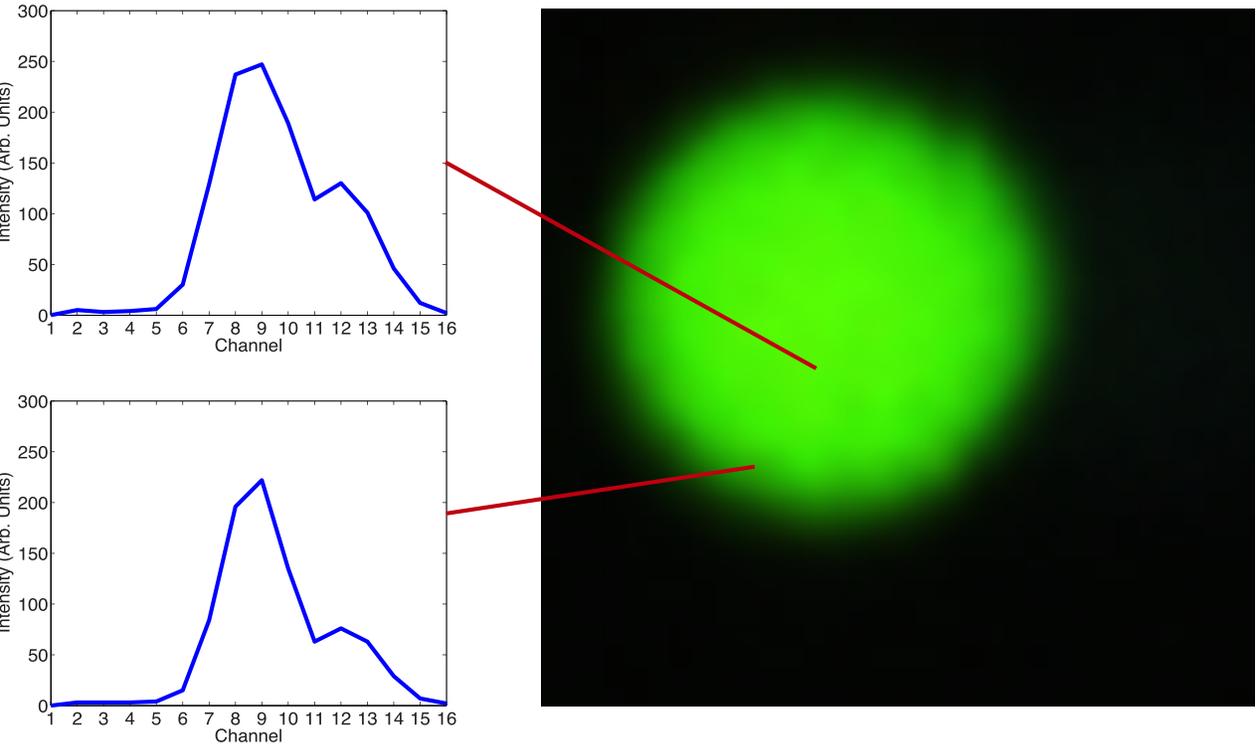


Figure 9b. Comparison of core vs. edge of FocalCheck bead: Green 2. The spectral content of the FocalCheck DoubleGreen bead 2 (525nm emission) is shown at the core and at the edge of the bead.

2" (525 nm). The third bead in **Figure 9c** is green 1 on the outside ring and green 2 in the center. Despite the spectral differences, all three false color images appear to have the same color.

The video in **Movie 3** shows a sonification of the same three image volumes. The cursor is clicked and dragged over each volume along a path that includes the blank space outside the bead, the edge of the bead, and the core of the bead. Choosing this path for each volume allows us to distinguish the "edge effect" of increasing intensity between the blank space and the bead (which is a function of scope resolution) from the subtle spectral shift between the thin ring around the outermost point of the bead and the bead's core.

Sonification of focalcheck beads

1 Data File

<http://dx.doi.org/10.6084/m9.figshare.4029513.v1>

Sonification of *Arabidopsis thaliana* specimen

The final test of our sonification was sonifying a real biological specimen that has been modified to express multiple fluorescent proteins. We imaged an *Arabidopsis* seedling sample that expressed three fluorescent proteins, RFP, YFP and GFP, which have been genetically modified in accordance to the procedure described by Kato *et al*²⁶.

In our first figure, we explore a visible similarity in the hyperspectral false color image between the plastids and the cell membranes of the *Arabidopsis* root. Although the plastids express a red fluorescent protein, background noise (likely from autofluorescence and scatter) in the green wavelength causes the plastids to appear very visually similar to the regions of the specimen expressing GFP.

However, when sonified, we can hear a very salient difference between plastids and the cellular membranes. Additionally, the plastids that are visibly red sound similar to the plastids with more green/yellow tinting, which allows the user to clearly understand the distribution of the plastids throughout the specimen despite the counterintuitive shift in color.

Sonification of Arabidopsis Plastid and Cell Membrane

1 Data File

<http://dx.doi.org/10.6084/m9.figshare.4029519.v1>

Next, we zoomed into an area of the specimen that has both mitochondria and cellular walls. The mitochondria express a protein (YFP) that is spectrally very similar to the protein the cell walls (GFP). In the false color image, this distinction is not apparent at all, due to the close proximity of the two spectral peaks.

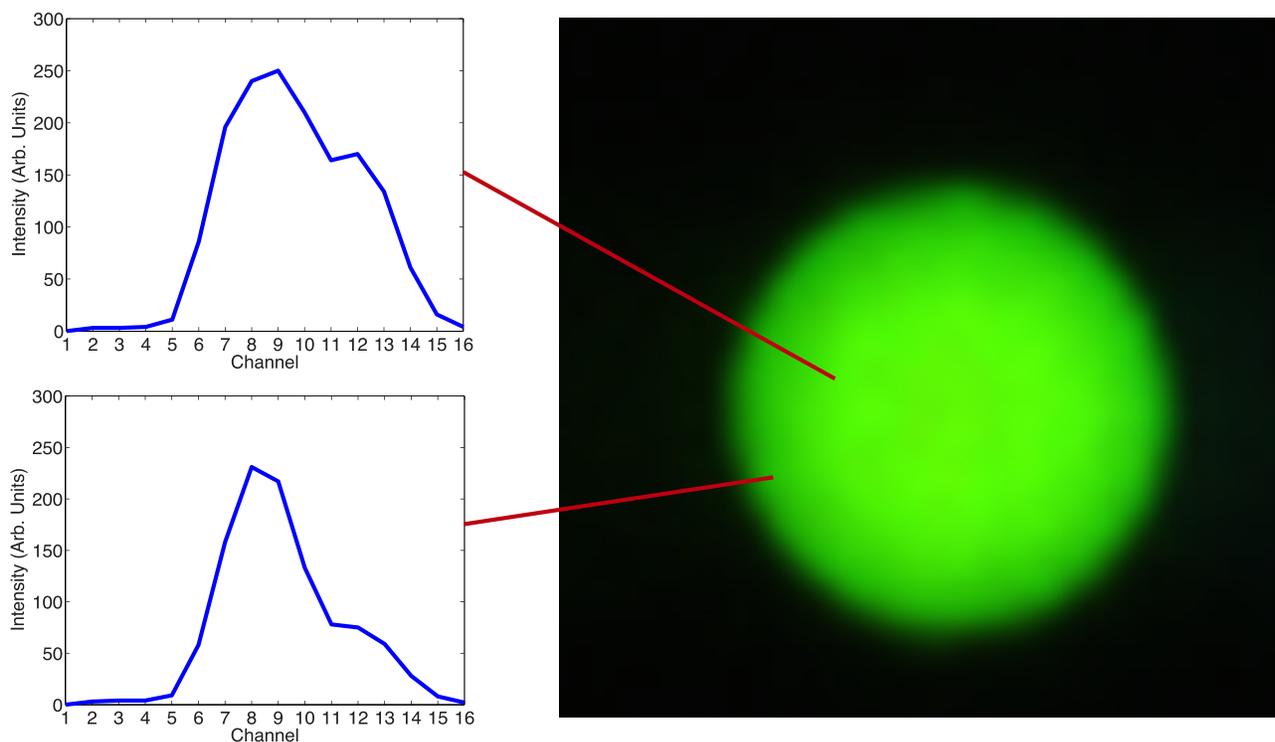


Figure 9c. Comparison of core vs. edge of FocalCheck bead: Mixed. The spectral content of the FocalCheck DoubleGreen bead 3 (512nm emission at the edge, 525nm emission at the core) is shown at the core and at the edge of the bead.

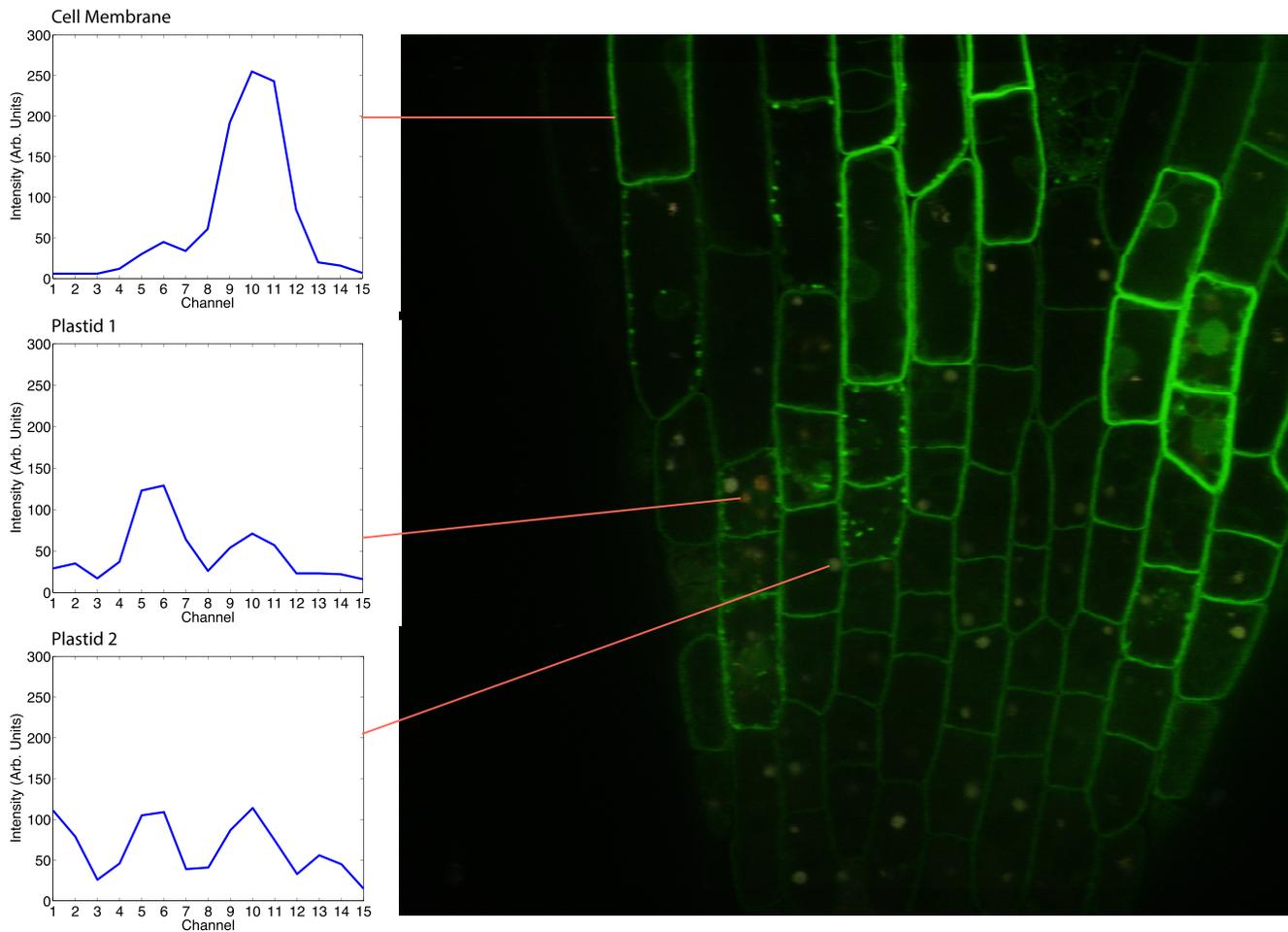


Figure 10. Spectral comparison of *Arabidopsis* plastid and cell membrane. This figure shows a hyperspectral false color image of an *Arabidopsis thaliana* specimen with spectral measurements of the raw multichannel image volume taken at various points of interest. The spectral differences between the cell membrane and the plastid are obscured in the hyperspectral image.

The sonification clearly articulates the difference between the two peaks, even at varying intensities. In this video a user sonifies several parts of the image where mitochondria are either very close to or overlapping the cellular boundary, and the spectral difference is still made apparent.

Sonification of *Arabidopsis* Mitochondria and Cell Membrane

1 Data File

<http://dx.doi.org/10.6084/m9.figshare.4029522.v1>

Discussion

Moving forward, the sonification framework that we have developed can be extended to provide deeper insights into complex microscopy datasets that have been inadequately represented through current visualization-only analysis methods. In particular, it is worth considering the potential utility of sonification-based

analysis in experiments that involve multiplexed functional markers, such as calcium signaling, or experiments that require differentiation between intrinsic and extrinsic fluorescence. Additionally, improvements to the functionality of our plugin would allow sonification based analysis to be utilized in more demanding contexts. Ideas for improvements to the plugin fall under two main categories: improvements to the sonification interface, and extension of the proof of concept.

Improving the sonification interface could be done on multiple levels that would amplify the utility of the plugin. First, adding functionality such as sonification over a Region of Interest, or automatic exporting of a sonification to a video format would afford the plugin a higher degree of similarity to traditional image analysis plugins, making it more usable to scientists already familiar with FIJI. On a deeper level, an external GUI could be designed to allow for full programmability of the sonification mapping process. Similar to the way that Digital Audio Workstations

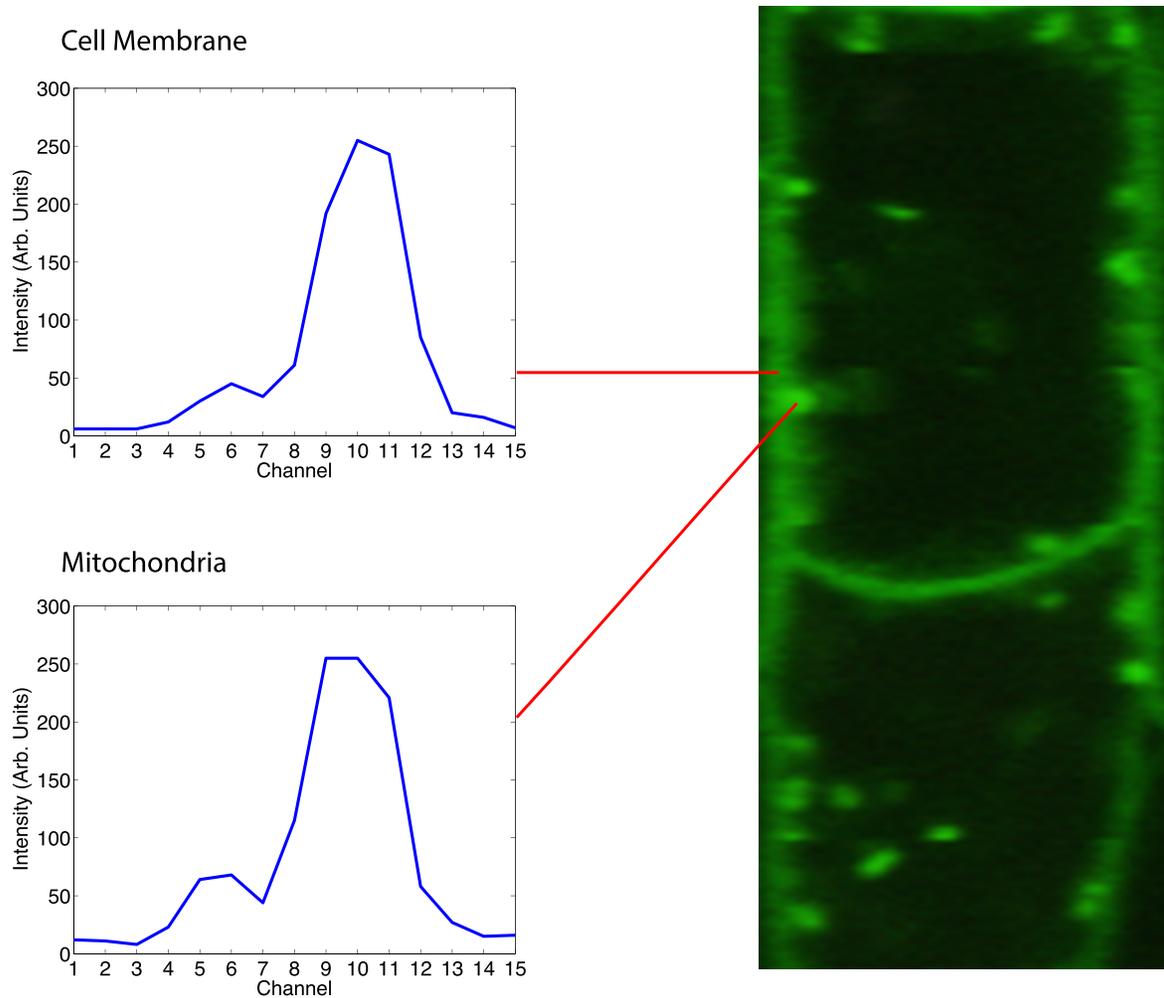


Figure 11. Spectral comparison of *Arabidopsis* mitochondria and cell membrane. This figure shows a zoomed in hyperspectral false color image of an *Arabidopsis thaliana* specimen with spectral measurements of the raw multichannel image volume taken at a cell membrane and a mitochondria. This presents another subtle spectral distinction in a live specimen that is not made clear by the hyperspectral image.

provide virtual instrument interfaces for programming custom sounds through a synthesis algorithm, our plugin could benefit from a visual interface that would allow users to tweak the mapping parameters connecting data space to sound space to create the perfect sonification for a given use case. This would be the audio equivalent of creating a custom lookup table for image analysis. Combined with the calibration image volume and the documentation of our incremental sound design process we have provided, users would be able to use such an interface to intuitively find the most useful sounds for given datasets.

Extending this proof of concept application into new use cases would also expand the functionality of our plugin. We demonstrate how sonification can be used to confirm a given computational spectral separation, and as such it would be worth considering the utility of sonification alongside other computational image analysis techniques utilized on spectrally dense datasets. Moreover, it is worth considering the utility of sonification as the size and

complexity of datasets scale, and whether parallel computing could potentially be useful for generation of sounds in real time with large numbers of input parameters. Finally, we have described ways that sonification provides us insights into our particular datasets that were inaccessible with only visualization, but it would be worth exploring in the general sense whether a difference can be quantified between a combination of sonification and visualization based analysis versus just visualization. Conducting an A/B test with a population of users and evaluating the variation in experience would be a good way to perceive the difference sonification makes in an applied context.

Conclusion

In this paper we have provided an in depth study of the utility of displaying spectral relationships in high-dimensional microscopy datasets. This has been the first study of sonification in microscopy without dimensionality reduction. Sonification has a wide variety of applications in image analysis, including providing data

display to visually impaired users, clearly conveying patterns and relationships in complex data, and providing a novel way to aesthetically represent a dataset. As multiple fluorescent markers are becoming more important and common in biological microscopy research and interpretation of the resulting rich spectral data become more relevant, scientists will have to turn to new types of tools to meaningfully interact with their data. Sonification has potential to not only be used as an alternative to conventional spectral visualization but as well to complement in certain use cases where sonification would be especially advantageous. We have presented sonification as more than just an aesthetic novelty, but as a promising solution to address limitations in existing spectral visualization based image analysis techniques. Future work will need to further explore the practicality of sonification of microscopy spectra in terms of appropriate sound mapping, biological application and user interface.

Data and software availability

F1000Research: Dataset 1. Raw data for ‘Sonification of hyperspectral fluorescence microscopy datasets’, Mysore *et al.* 2016, [10.5256/f1000research.9233.d138682](https://doi.org/10.5256/f1000research.9233.d138682)²⁷

Latest source code for the sonification plugin: <https://github.com/uw-loci/sonification>

Archived source code at the time of publication: [10.5281/zenodo.159960](https://doi.org/10.5281/zenodo.159960)²⁸

The plugin can be installed through the FIJI Updater by following the LOCI updates website (<http://loci.wisc.edu/software/fiji-plugins>). The calibration images utilized to test our synth designs are available on the LOCI GitHub alongside the plugin source code, so as to facilitate the development of new data-to-sound mappings by users.

Figshare: Video of Calibration. doi: [10.6084/m9.figshare.4029480.v1](https://doi.org/10.6084/m9.figshare.4029480.v1)²⁹

Figshare: Sonification of phantom image volume. doi: [10.6084/m9.figshare.4029495.v1](https://doi.org/10.6084/m9.figshare.4029495.v1)³⁰

Figshare: Sonification of focalcheck beads. doi: [10.6084/m9.figshare.4029513.v1](https://doi.org/10.6084/m9.figshare.4029513.v1)³¹

Figshare: Sonification of Arabidopsis Plastid and Cell Membrane. doi: [10.6084/m9.figshare.4029519.v1](https://doi.org/10.6084/m9.figshare.4029519.v1)³²

Figshare: Sonification of Arabidopsis Mitochondria and Cell Membrane. doi: [10.6084/m9.figshare.4029522.v1](https://doi.org/10.6084/m9.figshare.4029522.v1)³³

Author contributions

Aprameya Mysore implemented and applied the method generated artificial data, and wrote the paper. Andreas Velten collected real data, guided the development of data generation and sonification methods, and helped writing the paper. Kevin Eliceiri initiated and oversaw the project.

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no supporting grants were involved in this work.

Acknowledgements

The authors acknowledge technical support from the Laboratory for Optical and Computational Instrumentation (LOCI). The authors also thank Alexandra Chanoca and Marisa Otegui for providing the plants samples and Curtis Rueden for assistance with the FIJI implementation.

References

- Velten AU, White JG, Mackie TR, *et al.*: **Hyperspectral Multi-Point Confocal Microscope**. In *CLEO: 2013*. AW11.2. Optical Society of America, 2013. [Publisher Full Text](#)
- Lavis LD, Raines RT: **Bright ideas for chemical biology**. *ACS Chem Biol*. 2008; **3**(3): 142–155. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Pinaud F, Michalet X, Bentolila LA, *et al.*: **Advances in fluorescence imaging with quantum dot bio-probes**. *Biomaterials*. 2006; **27**(9): 1679–87. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Rizzo MA, Davidson MW, Piston DW: **Fluorescent protein tracking and detection: fluorescent protein structure and color variants**. *Cold Spring Harb Protoc*. 2009; **2009**(12): pdb.top63. [PubMed Abstract](#) | [Publisher Full Text](#)
- Zimmermann T: **Spectral imaging and linear unmixing in light microscopy**. *Adv Biochem Eng Biotechnol*. Springer, 2005; **95**: 245–265. [PubMed Abstract](#) | [Publisher Full Text](#)
- Conklin MW, Provenzano PP, Eliceiri KW, *et al.*: **Fluorescence lifetime imaging of endogenous fluorophores in histopathology sections reveals differences between normal and tumor epithelium in carcinoma *in situ* of the breast**. *Cell Biochem Biophys*. 2009; **53**(3): 145–157. [PubMed Abstract](#) | [Publisher Full Text](#)
- Conklin MW, Ada-Nguema A, Parsons M, *et al.*: **R-Ras regulates beta1-integrin trafficking via effects on membrane ruffling and endocytosis**. *BMC Cell Biol*. 2010; **11**: 14. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Sekar RB, Periasamy A: **Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations**. *J Cell Biol*. 2003; **160**(5): 629–633. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Harris AT: **Spectral mapping tools from the earth sciences applied to spectral microscopy data**. *Cytometry A*. 2006; **69**(8): 872–879. [PubMed Abstract](#) | [Publisher Full Text](#)
- Busch K, Howell N, Morrison GH: **Elimination of interferences in flame spectrometry using spectral stripping**. *Anal Chem*. 1974; **46**(14): 2074–2079. [Publisher Full Text](#)
- Zimmermann T, Rietdorf J, Pepperkok R: **Spectral imaging and its applications in live cell microscopy**. *FEBS Lett*. 2003; **546**(1): 87–92. [PubMed Abstract](#) | [Publisher Full Text](#)
- Davis LM, Shen G: **Extension of multidimensional microscopy to ultrasensitive applications with maximum-likelihood analysis**. In *Biomedical Optics (BIOS) 2007*. 64430N–64430N. International Society for Optics and Photonics, 2007. [Publisher Full Text](#)

13. Neher R, Neher E: **Optimizing imaging parameters for the separation of multiple labels in a fluorescence image.** *J Microsc.* 2004; 213(Pt 1): 46–62.
[PubMed Abstract](#) | [Publisher Full Text](#)
14. Pawley J: **Handbook of biological confocal microscopy.** Springer, 2010.
[Reference Source](#)
15. Kollmeier B, Brand T, Meyer B: **Perception of Speech and Sound.** In *Springer Handbook of Speech Processing.* Springer Science & Business Media, 2008; 65.
[Reference Source](#)
16. Olson HF: **Music, Physics and Engineering.** Dover Publications, 1967.
[Reference Source](#)
17. Hermann T, Ritter H: **Crystallization Sonification of High-dimensional Datasets.** *ACM Trans Appl Percept.* 2005; 2(4): 550–558.
[Publisher Full Text](#)
18. Nesbitt KV, Barrass S: **Evaluation of a multimodal sonification and visualisation of depth of market stock data.** 2002.
[Reference Source](#)
19. Lunn P, Hunt A: **Listening to the invisible: Sonification as a tool for astronomical discovery.** 2011.
[Reference Source](#)
20. Hermann T, Nattkemper T, Schubert W, *et al.*: **Sonification of Multi-Channel Image Data.** In *Proc of the Mathematical and Engineering Techniques in Medical and Biological Sciences (METMBS 2000).* (ed. Falavar, V.) CSREA Press, 2000; 745–750.
[Reference Source](#)
21. Cassidy RJ, Berger J, Lee K, *et al.*: **Auditory Display of Hyperspectral Colon Tissue Images Using Vocal Synthesis Models.** In *ICAD.* 2004.
[Reference Source](#)
22. Hermann T: **Sonification for Exploratory Data Analysis.** University of Bielefeld, 2002.
[Reference Source](#)
23. Schindelin J, Arganda-Carreras I, Frise E, *et al.*: **Fiji: an open-source platform for biological-image analysis.** *Nat Methods.* 2012; 9(7): 676–682.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Schneider CA, Rasband WS, Eliceiri KW: **NIH Image to ImageJ: 25 years of image analysis.** *Nat Meth.* 2012; 9(7): 671–675.
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Grond F, Berger J: **Parameter Mapping Sonification.** In *The Sonification Handbook.* Logos Publishing House, 2011; 363–397.
[Reference Source](#)
26. Kato N, Reynolds D, Brown ML, *et al.*: **Multidimensional fluorescence microscopy of multiple organelles in Arabidopsis seedlings.** *Plant Methods.* 2008; 4: 9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Mysore A, Velten A, Eliceiri K: **Dataset 1 in: Sonification of hyperspectral fluorescence microscopy datasets.** *F1000Research.* 2016.
[Data Source](#)
28. Rueden C, invisibility: **invizibility/sonification 1.0 [Data set].** *Zenodo.* 2016.
[Data Source](#)
29. Mysore A, Velten A, Eliceiri KW: **Video of Calibration.** *Figshare.* 2016.
[Data Source](#)
30. Mysore A, Velten A, Eliceiri KW: **Sonification of phantom image volume.** *Figshare.* 2016.
[Data Source](#)
31. Mysore A, Velten A, Eliceiri KW: **Sonification of focalcheck beads.** *Figshare.* 2016.
[Data Source](#)
32. Mysore A, Velten A, Eliceiri KW: **Sonification of Arabidopsis Plastid and Cell Membrane.** *Figshare.* 2016.
[Data Source](#)
33. Mysore A, Velten A, Eliceiri KW: **Sonification of Arabidopsis Mitochondria and Cell Membrane.** *Figshare.* 2016.
[Data Source](#)

Open Peer Review

Current Referee Status:



Version 1

Referee Report 24 November 2016

doi:10.5256/f1000research.9937.r17182



Andrew R. Cohen

Department of Electrical and Computer Engineering, Drexel University, Philadelphia, PA, USA

Review for: Sonification of hyperspectral fluorescence microscopy datasets

Summary:

This paper by Mysore, *et al.*, describes a framework and ImageJ plugin for constructing an audio signal from, or sonifying, 2-D multichannel fluorescence microscopy images. There are many positive aspects to this work. There is a pressing need for more effective approaches for human interaction with large, high-dimensional datasets. The use of audio information to enhance multi-channel image data is a key aspect of effective visualization that has not been well investigated. This is an important research topic. The manuscript is clearly written. The software is free and open source, with a broadly applicable ImageJ plugin provided. These are all commendable aspects of the manuscript.

There are however, a few shortcomings in both the approach and the manuscript that should be addressed in the next revision. Most significant, the approach proposed here seems to suggest that sonification is an alternative to visualization, rather than an enhancement to visualization. It seems clear that sound will greatly enhance visualization, but the manuscript states in multiple places that sound is superior to visualization. The result is the implication that sonification is an alternative rather than an enhancement to visualization. This is a surprising and seemingly unsupported notion. This point needs clarification.

Specific Comments:

1. The authors repeatedly refer to the discriminative limitations of the human visual system compared to auditory processing. For example, "While the human vision system is limited for natively comprehending rich hyperspectral data, the human ear and auditory system are more optimally suited for this task¹". This statement is untrue, and it is not supported by this reference that makes no relative measure in discriminative ability between human audio vs. visual capacity:
2. Related comment: "whereas the eye has only three channels of spectral sensitivity, the ear can distinguish about 1400 pitches throughout its range" is another misleading statement about human visual capacity. The visual system is generally superior to the auditory system at subtle discrimination tasks. The human eye is remarkably sensitive to color variations. If visual data has only three channels, then audio has at best two (stereo) and in the present approach just one channel. One key aspect of visual vs. auditory capabilities is the color map (or tonal map) used to render image (sound) data.... Please comment on the established literature quantifying relative

bandwidth and discriminatory capabilities of the human visual vs. audio perception.

3. ...leading to another related point: “False color images of datasets with three or more spectral dimensions can visually appear to be the same in areas where the data are vastly different, because the transcoding operation used to render the raw data visible to human eyes involves a data loss”. Color mapped rendering is only lossy when the number of colors used exceeds the size of the color map. How many unique intensity bins come out of the microscope? Often, 8 bit per pixel images are sufficient quantization, and at this level it is straightforward to design a colormap that not only preserves intensity differences, but that enhances them in a perceptual manner. Even for 16 bit color images it can be possible to design effective color mappings. see e.g.²
4. The introduction refers to spectral unmixing as a bottleneck in multi-channel image analysis, but does not clarify why or how audio should be used as part of the solution to the spectral unmixing question. Put another way, how is the mouse-driven sonification as described in the present paper used for spectral unmixing?
5. What is the source for Figure 1? Figure 2 seems to be sourced from references^{1,3} – is this correct? Or was this somehow measured as part of the present work?
6. Dimensionality reduction is only a concern if meaningful information is lost. Ultimately, all biological image analysis involves dimensionality reduction – from the raw image dataset to a pie chart or distribution or some such representation. The manuscript states that “By allowing researchers to interact directly with the spectral composition of their datasets, we can potentially elucidate salient characteristics of the data that otherwise would have gone unnoticed.” Do you have an example of this for a real-world dataset? The example datasets (Figures 9-11) all seem to show data that is not clearly visually discriminated, but this seems a function of poor visualization rather than inherent visual limitation (see more detailed comments on color mapping above). Are there differences that can be detected via sonification that would not be detectable by a human using *effective* visualization, or by a computer using statistical pattern recognition techniques?
7. How about feature-driven sonification rather than pixel driven? The manuscript alludes to this in stating that “since FIJI is used for many spectral fluorescence microscopy analysis routines, it opens up the possibility of integrating and interacting with these steps as needed by the end user.” This seems a much more logical and promising approach – why isn’t it explored in more detail here?
8. Does the method work with 3-D multi-channel images? If so, please explain how. If not, then this seems an important limitation that should be explicitly acknowledged.
9. The system is driven by clicking and dragging the cursor across the image. This seems a cumbersome method, and a good illustration of the lack of spatial information in the audio signal. This also seems to introduce a new bottleneck into the analysis pipeline, particularly as the image resolutions continue to increase, e.g. to 4k resolutions and beyond. Consider this in light of the 450x450 pixel sample images. Can you comment on the scalability of the approach?
10. The sound design methodology fails to account for a perceptually “pleasing” sound scheme. The present tonal model becomes quite annoying after a short amount of interaction with the image data – is there a way to capture a less abrasive representation?

11. The manuscript states that “Another issue related to context we observed was that, similar to the case with color lookup tables, there is no “one size fits all” auditory display solution”. There are in fact a number of generally effective color mapping approaches. Specifically, there exists a significant amount of research on quantitative color spaces (e.g. HSV, perceptual spaces as mentioned above, etc.). Are there equivalent theoretical bases for audio? If not, why not? If so, please reference.
12. Figure 6 seems misleading. Clearly a color map that can more effectively combine the channels could be formulated, particularly given the interactive environment that is enabled by sliding a mouse across the image for an audio comparison.
13. Figure 8 also seems contrived – clearly it is possible to construct a multi-channel image that naively combines to any monotone. At the same time, it would be straightforward to e.g. learn a color space that maximizes separation among the image channels just from the given data.
14. What would Figure 10 look like with an HSV color map? Just because the fluorophores are colored green and red does not mean that is the best color choice for representing them in a visualization context.
15. Following the questions of naively designed color maps being easily defeated, are there equivalent limitations in your audio mapping that would allow images that are visually different to sound identical? Could you construct such an example?
16. What about other senses in addition to sound? E.g. touch has been explored in data visualization previously⁴ – would such approaches add value in combination with your method?

References

1. Kollmeier B, Brand T, Meyer B: Perception of Speech and Sound. *Springer Handbook of Speech Processing, Springer Science and Business Media*. 2008; **65**.
2. Schanda J: Colorimetry: Understanding the CIE System. *John Wiley and Sons*. 2007.
3. Olson HF: Music, Physics and Engineering. *Dover Publications*. 1967.
4. Koerner M, Wait E, Winter M, Bjornsson C, Kokovay E, Wang Y, Goderie SK, Temple S, Cohen AR: Multisensory interface for 5D stem cell image volumes. *Conf Proc IEEE Eng Med Biol Soc*. 2014; **2014**: 1178-81 [PubMed Abstract](#) | [Publisher Full Text](#)

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 23 November 2016

doi:[10.5256/f1000research.9937.r17919](https://doi.org/10.5256/f1000research.9937.r17919)



Paul S. Maddox

Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

In the manuscript titled “Sonification of hyperspectral fluorescence microscopy datasets”, Mysore, Velten and Eliceiri present the development of an “eye-to-ear” conversion plugin for Fiji. Modern imaging places a premium on extracting as much information as possible from images and much focus has been placed on increasing the dimensionality, and thus ability to segment and measure, of image data. These efforts have resulted in images containing 6 or more dimensions (x, y, z, Intensity, time, color, fluor lifetime, etc), making presentation and interpretation visually complex. In the case of spectral imaging, commercial instruments are fully capable of producing 32 channels in the color range (hyperspectral), further complicating the situation. Here, the authors present an innovative solution: a D to S conversion (digital to sound).

The human eye is tuned to respond to a relatively narrow band in the electromagnetic spectrum, however to do so with relatively high resolution. On the otherhand, the human ear has a very high (~3 orders of magnitude) dynamic range combined with high resolution. As such, the ear is better suited to interpreting hyperspectral data. The authors make this case effectively with supportive calibration and experimental data. This is an exciting window (or tune?) into the future of data presentation.

Below I make a few suggestions:

1. The authors do an excellent job of explaining the biology of the eye, however the ear seems to be less well defined. What about the biology, mechanics, or innervation of the ear allows this dynamic range? Personally, I have always considered the ear a single dimensional detector; at any given moment, the ear interprets and reports a single piece of data. Is this the case or can the ear send multiple inputs to the audio processor in our brain?
2. Can this concept be applied to non-spectral information? Have the authors tried to use it for simply portraying depth or time? In my experience there are only a limited number of people using spectral data when compared to depth and/or time.

Minor comment:

Page 3; what is “image UI”?

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.
