

HHS Public Access

Int J Comput Assist Radiol Surg. Author manuscript; available in PMC 2019 January 01.

Published in final edited form as:

Author manuscript

Int J Comput Assist Radiol Surg. 2018 January ; 13(1): 25-35. doi:10.1007/s11548-017-1667-5.

Auditory display for fluorescence-guided open brain tumor surgery

David Black,

Medical Image Computing, University of Bremen; Jacobs University, Bremen; Fraunhofer MEVIS, Bremen, Germany

Horst K Hahn,

Fraunhofer MEVIS, Bremen, Germany; Jacobs University, Bremen, Germany

Ron Kikinis,

Medical Image Computing, University of Bremen; Fraunhofer MEVIS, Bremen, Germany; Brigham and Women's Hospital and Harvard Medical School, Boston, USA

Karin Wårdell, and

Department of Biomedical Engineering, Linköping University, Linköping, Sweden

Neda Haj-Hosseini

Department of Biomedical Engineering, Linköping University, Linköping, Sweden

Abstract

Purpose—Protoporphyrin (PpIX) fluorescence allows discrimination of tumor and normal brain tissue during neurosurgery. A handheld fluorescence (HHF) probe can be used for spectroscopic measurement of 5-ALA-induced PpIX to enable objective detection compared to visual evaluation of fluorescence. However, current technology requires that the surgeon either views the measured values on a screen or employs an assistant to verbally relay the values. An auditory feedback system was developed and evaluated for communicating measured fluorescence intensity values directly to the surgeon.

Methods—The auditory display was programmed to map the values measured by the HHF probe to the playback of tones that represented three fluorescence intensity ranges and one error signal. Ten persons with no previous knowledge of the application took part in a laboratory evaluation. After a brief training period, participants performed measurements on a tray of 96 wells of liquid fluorescence phantom and verbally stated the perceived measurement values for each well. The latency and accuracy of the participants' verbal responses were recorded. The long-term memorization of sound function was evaluated in a second set of 10 participants 2–3 and 712 days after training.

Compliance with ethical standards

Conflict of interest

The authors state that they have no conflict of interest.

Ethical Approval

No ethical approval is required, as the study was not a type that could affect the subjects physically or psychologically. Data used in Fig. 1 are from a study with ethical approval described in [5].

Informed consent

Informed consent was obtained from all individual participants included in the study.

Results—The participants identified the played tone accurately for 98% of measurements after training. The median response time to verbally identify the played tones was 2 pulses. No correlation was found between the latency and accuracy of the responses, and no significant correlation with the musical proficiency of the participants was observed on the function responses. Responses for the memory test were 100% accurate.

Conclusion—The employed auditory display was shown to be intuitive, easy to learn and remember, fast to recognize, and accurate in providing users with measurements of fluorescence intensity or error signal. The results of this work establish a basis for implementing and further evaluating auditory displays in clinical scenarios involving fluorescence guidance and other areas for which categorized auditory display could be useful.

Keywords

fluorescence-guided resection (FGR); spectroscopy; 5-aminolevulinic acid (5-ALA); protoporphyrin (PpIX); surgical navigation; neurosurgery; human-computer interaction; user interfaces; sonification; LabVIEW

1 Introduction

Fluorescence imaging based on 5-aminolevulinic acid (5-ALA) visualized using fluorescence-guided resection (FGR) surgical microscopes is an optical guidance system that has been introduced during the past decade for routine clinical application [1]. 5-ALA is a photosensitizer administered prior to the surgical procedure, which is metabolized and accumulated as protoporphyrin (PpIX) in the tumor cells. When excited by light, the tumor re-emits fluorescence of PpIX, thereby enhancing visibility to the surgeon.

The spectroscopic measurement technique applied using a fiber optic probe is based on measurement of a fluorescence spectrum, which is usually displayed on a screen in the OR. A handheld fluorescence (HHF) probe for spectroscopic measurement techniques has been developed at Linköping University. This HHF probe has been evaluated in over 50 patients in the OR for brain tumor resection guidance as a stand-alone system [2,3], and in combination with both a neuronavigation system [4] and an FGR microscope [5]. When operating with the FGR microscope, the surgeon observes both the surgical site and the fluorescence without requiring additional feedback support. The fluorescence seen through the FGR microscope is conventionally grouped into negative, weak, and strong [1,5]. This categorization supports the neurosurgeon in decision making on tissue removal. However, for HHF probe measurements, the surgeon cannot reliably perceive the fluorescence signals, specifically weak signals, through vision alone. Moreover, several complications during the operation can disturb the measurement, including blood interference and, to a lesser extent, a surgical microscopes white light lamp or system failure, including misplacement of the probe. In these cases, the surgeon should be informed to correct for the measurement error. Currently, intraoperative fluorescence measurements from the HHF probe are provided on a computer screen and interpreted by an engineer responsible for the system, who verbally relays signal values in the OR (Fig. 1a). Examples of fluorescence signals measured during surgery are shown in Fig. 1b. Ideally, the surgeon should be able to receive information from the system without the need for an interpreter while keeping the visual focus on the surgical

site. The responsibility of signal interpretation cannot be placed on the surgeons or their assistants; therefore, additional visual or auditory support for acknowledging measurement results could facilitate this process.

Utsuki et al. reported a system which triggered an audible tone when the measured PpIX fluorescence intensity between 632 and 636 nm exceeded a certain level. Neither the sound characteristics nor intensity varied with the changes in fluorescence intensity and did not consider errors that occur during intraoperative measurements [6]. Further reports on audible systems for optical measurements have been limited. To the authors' knowledge, this paper presents the first and only evaluation of auditory feedback to relay fluorescence intensity values of an HHF probe.

Using sound to transmit changes in data, termed auditory display, has recently gained attention for a small but varied array of clinical applications to aid surgeons during image guidance. Examples in the literature describe auditory display as a means of information retrieval that goes beyond monitoring tasks (such as those used in anesthesia [7]) and helps deliver important navigation information to the clinician to reduce reliance on computer screens or to enhance awareness of important anatomical risk structures in the vicinity of the surgical instrument. Previous implementations of auditory display for image-guided interventions include neurosurgical volume resection [8,9], temporal bone drilling [10], cochlear implantation [11], liver resection path marking [12], and ablation and biopsy needle placement [13,14]. Advantages found in previous attempts include heightened awareness of or distance to anatomical risk structures [10,11,15], reduced surgical complication rate [15], increased visual focus on the surgical site [12], and improved placement accuracy [14] when using auditory display as either to replace or to augment existing visual support systems. For a review of applications of auditory display in image-guided interventions, see [16].

The aim of this study was to develop and evaluate an auditory display to support fluorescence-guided open brain tumor surgery using an HHF probe in the laboratory based on previously determined clinical fluorescence intensity levels. The investigation is a first evaluation of auditory display to support HHF probe fluorescence measurements and can be implemented into similar optical systems. By using auditory display to support fluorescenceguided brain tumor surgery, the surgeon should be able to "hear" fluorescence values without relying on a computer screen, thereby enhancing visual focus on the surgical site. In addition, auditory feedback should reduce the need for a surgical assistant to verbally relay intensity values to the surgeon, thus receiving values more quickly and reducing missed values due to interpersonal miscommunication.

2 Methods

2.1 Experimental design

The principles of intraoperative fluorescence measurements and fluorescence quantification are described in previous work [2,3,5]. The fluorescence signal levels in this study were replicated in a tray of optical phantoms (Fig. 2a) to be comparable to the levels measured in the OR [2,5]. The instrumentation was selected for the study so that all hardware

components were compatible with the LabVIEW software (National Instruments, Inc., Austin, TX) and Open Sound Control (OSC) [17] protocol.

2.1.1 Brain tumor phantoms—Four sets of liquid phantoms were prepared to model the actual clinical measurement situation (see Table 1). These included zero signal, low signal, high signal, and error signal. The phantoms modeled the optical properties of the brain tumor using ink and intralipid 20% (Fresenius Kabi, Uppsala, Sweden) [18] including tissue autofluorescence (AF) by adding turmeric dissolved in ethanol (zero signal). In two phantom sets (low and high signal), 10 and 30 g/l of PpIX disodium salt (MP Biomedicals, France) was added to model the low and high fluorescence signals, respectively. The PpIX concentration was chosen to be greater than what is measured in the brain to account for photobleaching effects on the signals and thus avoid variation in the generated sound on one spot. The maximum PpIX peak in the phantoms was at 634 ± 4 nm due to the chemical environment. In a fourth phantom set, the AF was blocked by additional ink to reflect the situation in which the measurements are obstructed by blood or no signal is recorded (error signal). The tray had 96 wells each of 7 mm diameter and 1 cm depth, see Figure 2a.

2.1.2 Hardware and signal analysis setup—A 405-nm laser (Oxxius) in continuous mode was used to excite the fluorescence, together with an AvaSpec-ULS2048L-USB2 spectrometer (Avantes BV, Netherlands) for detection of fluorescence, measuring wavelengths 580–1100 nm. A fiber optic probe (Avantes BV, Netherlands) was used to measure the fluorescence in the phantoms. The probe included one central fiber for excitation and six surrounding fibers for fluorescence collection. The total diameter of the fiber bundle was 1.2 mm. A custom spectrometer interface was developed in LabVIEW. The OSC library was embedded in the program to send measurement values to the sound synthesizer. The spectrometer integration time was set to 800 ms to reflect the settings used in the OR.

The intensity from the 600 nm wavelength representing AF and 630 nm wavelength representing PpIX fluorescence was extracted for the analysis. The intensity of the AF at 600 nm was first analyzed to determine whether an error was present. If the intensity was lower than a certain threshold, a signal level was generated at a very large value out of the fluorescence range, in this case 5000. The threshold was set to the intensity of AF at 600 nm measured on the error phantoms after phantom preparation and the average on the zero phantoms. If the intensity was lower than the threshold, the error tone was generated. If the intensity was higher than the upper threshold, the value was set to a constant within the thresholds' range. This loop was added to compensate for the effect of AF from the well side walls. If no error was identified, quotient of fluorescence intensity at 630 nm and 600 nm was calculated and sent as a "pulse" to the sound synthesizer. The set of threshold quotients for the conversion of these values to the various tones are included in Table 1. The principle of fluorescence signal conversion to sound is shown as a flowchart in Fig. 3. The program required that the values first be calculated after the signal was measured, creating a delay of one pulse. The tone was played back instantly after a new intensity value was sent to the synthesizer.

2.2 Auditory display design

An initial experimental study was conducted with one neurosurgeon well acquainted with the measurement system and clinical application. Various synthesis methods were presented to map the intensity of the fluorescence signal to parameters of the auditory display. These included a continuous mapping of intensity to vibrato (frequency modulation) rates, continuous mapping of intensity to the pitches of two alternating tones for comparison, and mapping intensity into discrete values for the playback of individual, short composed musical note sequences to represent desired functions, so-called earcons [19]. After the initial study, discrete values were selected as the optimum mapping method, as the categorization of intensity mapped to a selection of a small number of earcons was found to be most appropriate for the clinical task. Continuous auditory displays, while beneficial for auditory display for surgical trajectory navigation in general [12], are harder to translate into quantitative values, in which case a classification-based approach is suggested. This has, for instance, been successfully employed for auditory display for awareness of risk margins in image-guided interventions [11]. Thus, a set of four tones were produced which transmitted one of three intensity levels (zero, low, and high signal) or the error signal.

The following four tones were produced:

Zero signal The signal was generated when the fluorescence intensity quotient was 0– 0.9. The zero signal tone informs the user that the signals were measured correctly but no PpIX fluorescence (tumor) was detected. This was synthesized as a cluster of three sine wave generators with frequencies of 233, 277, and 370 Hz which were played back with an amplitude envelope of 200 ms attack phase, 400 ms sustain phase, and 200 ms release phase. The resulting tone was a calm major G chord ¹ with a total time of 800 ms, see Figure 4a.

Low signal The low-signal tone played back when intensity quotients ranged from 1 to 12.9. The tone consists of two consecutive triangle wave pulses with frequencies of 220 and 260 Hz. The amplitude envelope consisted of a 30 ms attack phase, 20 ms sustain phase, and 300 ms release phase. The 260 Hz pulse played back 233 ms after the start of the 200 Hz pulse. The resulting tone is more intense than the zero-signal tone, as the attack time was shorter, there were more pulses played back, and the triangle wave contained a higher number of upper-level harmonics (resulting in increased brightness) than the sine pulses of the zero-signal tone; see Fig. 4b.

High signal The high-signal tone played back when intensity quotients ranged from 13 to 50. The tone consisted of four sequential triangle waves with frequencies of 349, 440, 523, and 698 Hz. The amplitude envelope of each triangle wave pulse consisted of a 30 ms attack phase, 20 ms sustain phase, and 300 ms release phase. The sequential triangle waves played back with inter-onset intervals of 70 ms. The resulting tone was even more intense than the low-signal tone, indicating an increased

¹Eighteenth-century composer Christian Schubart described the major G chord as evoking "everything rustic, idyllic and lyrical, every calm and satisfied passion, every tender gratitude for true friendship and faithful love - in a word, every gentle and peaceful emotion of the heart is correctly expressed by this key" [20]

Int J Comput Assist Radiol Surg. Author manuscript; available in PMC 2019 January 01.

urgency due to a higher number of upper-level harmonics and shorter times between sequential pulses; see Fig. 4c.

Error Finally, the error tone consisted of two simultaneous triangle wave pulses with frequencies of 500 and 515 Hz played back three times with a delay of 100 ms. The amplitude envelope of each pulse consisted of a 20 ms attack phase, 50 ms sustain phase, and 20 ms release phase; see Fig. 4d.

Thus, an auditory display synthesizer was created to play back three intensity tones and one error tone. The intensity tones from zero to low to high featured increasing frequency, number of pulses, and onset speed. The four tones were synthesized and played back using the Pure Data [21] sound synthesis environment in real time. Intensity levels were sent at an interval of 800 ms from the fluorescence measurement system to a computer hosting the sound synthesis environment. Playback used a pair of standard multimedia loudspeakers connected to the synthesis computer located approximately 1 m in front of the user.

2.3 Method evaluation

2.3.1 Participants—Participants (n = 20; male= 10 and female = 10) ranged from ages 23 to 56 (median 25) years old, and none self-reported hearing or vision impairment. The first set of 10 participants were included in the function response test, and the second set of 10 participants were included in the memory test. Self-response proficiency in music of the first group of participants included professional (n = 1), amateur (n = 4) and no musical (n = 5) experience, and in the second group professional (n=1), amateur (n = 5) and no musical experience (n=4). The majority of the participants were chosen from a population with backgrounds in engineering or natural sciences but without any professional experience with optical measurements, auditory display, or the project's background.

2.3.2 Experimental setup—The laboratory experiment setup was designed to assess the ability of participants to distinguish fluorescence readings between three intensity levels and an error signal using the auditory display. This was produced to mimic the ideal situation in the OR, where the surgeon should be informed of intensity levels and error states without an assistant. The experiment consisted of a preliminary intuition test which asked the participants to guess which auditory display tone corresponds to which fluorescence measurement function, and a subsequent function response test in which the participants were asked to measure the fluorescence of a series of phantom wells using the HHF probe and state which function was played back by the auditory display.

2.3.3 Intuition test—For the intuition test, each participant assigned the 4 played tones to functions, including zero signal, low signal, high signal, and error, resulting in 4 data points per participant and 40 data points in total for all participants. To avoid biasing, participants were provided with no information about the application and principles of the measurements. The four tones were played back in a fixed, randomized [22] order used for all participants. First, the tones were played back with each tone repeated three times. Then, the tones were played back again, this time only once. Thereafter, the tones were played back once again, and participants were asked which tones they would assign to each

function (zero, low, high or error signal). A final round of playback was undertaken so that participants could change their answer if desired.

2.3.4 Function response test—During the function response test, participants navigated the HHF probe across a tray of phantom wells (Fig. 2a) and verbally stated the function perceived for each well. As a training task, participants used a tray orientation that differed from that used later in the test. During training, participants could ask for help or start or stop navigation at will. To complete the training, participants moved the HHF probe over the wells and stated the perceived tone function after hearing two consecutive, equal readings. The training was performed on 32 training wells, after which all participants confirmed having successfully become familiar with the tones. No response time or accuracy data were recorded during the training phase.

After the training phase, including a 1-min pause, the test procedure was performed for each participant using a 96-well tray (8 rows, 12 columns); see Fig. 2b. During the test procedure, navigation using the tray was divided into 3 sets of 4 columns each, with a 1-min pause between sets. The actual value and the verbally stated value for each well and each participant were recorded on video.

The tray was filled with phantoms evenly prepared in clusters of 4 values which were arranged across the tray. The sequence of the 24 permutations of the 4 played back tones was randomized [22], and the same sequence was used for all participants; see Fig. 2b. The phantoms were mixed with an even distribution throughout the well to test all the sequences of tones hearing between zero, low, high, and error signal; see Fig. 2b. Each well was visually indistinguishable.

To perform the function response task, participants navigated through the wells by moving the tray so that one well was situated directly beneath the probe. The navigation sequence was such that participants measured all wells in one column and then moved on to the next column. After the fluorescence was measured by the HHF probe, tone was played back through the loudspeakers. As in the training phase, participants were instructed to only state the perceived function after being confident of its stability by listening for two consecutive, equal tones using the auditory display. After stating the perceived function ("zero," "low," "high," or "error") the participant moved the tray so that the next probe measurement could be taken. This process was repeated until all wells in the tray had been measured and the perceived functions stated. The subjects were not provided with any visual clues of the measured signal while evaluating the tones.

2.3.5 Memory Test—The intuition test was performed at the first step. The participants were then instructed on the intended function and trained by playing the tones in a random order [22], after which the participants gave their response. This was repeated for approximately five rounds of playback. The memory was tested on days 1–3 and days 7–12 depending on the participant availability. For the memory test, the tones were played in a random order once and afterward one by one in the same order when the participant responded.

2.4 Data analysis and statistics

For each phantom well, the played sound value was compared to the participants response. Using power sample size calculation, a minimum of 55 samples were needed to achieve 98% accuracy with a power of 0.95. The number of measurements for each participant was approximately twice as much as this value. In total, 960 data points were recorded. The total number of played tones for each function were 255 zero signals, 263 low signals, 269 high signals, and 173 error signals. No data were excluded from the analysis.

Accuracy was calculated as the ratio of the total number of correctly identified tones to the total number of played tones for each tone and each participant. The latency of the response was calculated by the number of pulses (single, played back tone) needed until the participant uttered the response. The number of pulses was recorded, including the minimum two initial pulses, that the participants needed before uttering a response. Pulses uttered during playback of a tone were recorded as having the previous number of pulses. For instance, if a participant responded while the fourth pulse played back, this was recorded as a latency of 3 pulses.

The statistical tests were performed in MATLAB R2015a (MathWorksTM, Inc.). The null hypothesis was that the played sounds were not distinguishable and that there was no correlation between the played and perceived sounds. As some of the datasets were not normally distributed, the Mann-Whitney test was used for assessing statistically significant difference, where p < 0.05 was considered to show a statistical significance. Linear correlation was used for assessing the goodness of fit (R^2) between each of the two datasets. Boxplots were used to represent the data sets where the mid-line in the box was the median and the box was set to 25–75% quartiles.

3 Results

3.1 Intuition test

Twenty-one of the 40 reported tone intuition assignments corresponded to the intended function, resulting in a true-to-total ratio of 52%; see Table 2. Tone two was the most intuitive, as all participants correctly assigned this to the high-signal function. Tones zero and one were the least intuitive; 50% assigned the zero signal tone to low-signal function, and 60% assigned the low-signal sound to the zero-signal function. For the error tone, 60% correctly assigned the correct signal, whereas 20% assigned this to the zero-signal function and 20% to the low signal function.

3.2 Function response test

3.2.1 Response accuracy—The training period for the 10 participants lasted an average of 80.44 (\pm 29.44) s. After training, 943/960 (98%) of the participants responses matched the played function. Recorded values versus verbally given responses are shown as a confusion matrix in Table 3. The highest confusion was exhibited for the zero-signal tone: It was perceived in 3% of the cases to be the low-signal tone. The high-signal tone was the most confidently perceived, as more than 99% of responses were correct. There was no statistically significant difference between the played tone and the responses for any of the

participants when each played tone was compared to each verbal response for each participant (*p*-value for each participant > 0.8). There was a low and negative correlation between the median accuracy and the musical proficiency for each participant ($R^2 = 0.3$), but no correlation was found between the accuracy in the initial intuition test before the training and the accuracy after training ($R^2 < 0.1$).

3.2.2 Response latency—The response latency (time needed to respond verbally) was a minimum of 2 pulses (1.6 s), and at most 9 pulses (7.2 s) with a median of 2 (average = 2.6) pulses across all participants. The response latency for each participant is plotted in Fig. 5. When the latency for each tone was separately analyzed for each participant, only in participant no. 6 was the median different for separate tones; therefore, the latency showed a dependence on the individual rather than on the tone. The correlation between the median latency and accuracy ($R^2 < 0.1$) for each participant was negligible (Fig. 6). The musical proficiency (0, 1, 2) did not show any correlation with the median response latency for each participant ($R^2 = 0$).

3.2.3 Memory response—The initial intuition test for this population was 60% accurate. After becoming acquainted with the tone functions, the responses during the training were 100% accurate both on days 2–3 and days 7–12.

4 Discussion

The goal of this study was to investigate the practicality of using auditory displays for the communication of PpIX measurement results based on those used during neurosurgery. The experimental study was designed to determine how well the intensity of a fluorescence measurement using an HHF probe could be recognized by listening to tones played back in a laboratory environment. The auditory display for fluorescence intensity was developed to provide the surgeon with four discrete tones to help identify the status of the fluorescence measurement without having to view a computer monitor or rely on the support of an assistant to verbally relay intensity values. The principles can be applied to any other PpIX fluorescence spectral analysis algorithm; however, the limits of zero, low, and high signals would depend on the device and the analysis algorithm [23–26].

4.1 Function response

The participants intuitively assigned tone functions were evaluated, as well as how quickly and accurately participants could respond to the played tones while measuring a tray of fluorescence phantoms. The results of the evaluation show that in almost all cases, participants correctly identified each of the four played intensity tones after a median of 2 pulses (1.6 s), where two pulses was the minimum amount of time to correctly provide an intensity measurement. The intuition test showed that even before using the system or receiving training, participants were able to correctly assign the synthesized sounds to the intended functions in more than half of all cases. After training, the response accuracy increased to 98% for the function test, and the memory test was 100% accurate for all participants after periods of both 2–3 days and 7–12. Thus, the developed auditory display is easy to learn and remember, quick to use, and highly accurate. Even in its current form, the

auditory display could provide an immediate benefit to surgeons by delivering intensity values without the need for a screen or reliance on a surgical assistant to relay values.

The participants were not provided with any information on the principles of fluorescence or the intended purpose of the measurements to reduce the effect of being assisted by vision. The unintentional visual assistance which might have been induced in some cases at the phantoms with low and high PpIX concentration was considered negligible as the participants could not interpret the colors and the colors were blurry without any optical filter. Measurement accuracy had a median of 66% (range 35–95%) weakening any chances of visual assistance. This measurement error which was caused due to the wall of the phantom wells does not occur in the actual situation during operation.

4.2 Auditory display design

The auditory display should be designed to be easy to learn, and tones should be distinguishable from one another, so that the surgeon is able to understand the meaning of the display in a few seconds. The auditory display should be fast enough to transmit the desired intensity to the surgeon within the duty cycle of the system, i.e., the playback of the entire tone should occur between successive measurements. The results of the evaluation show that on average, participants could recognize the sound and verbally respond with the correct value directly within the duty cycle of the system even after undergoing only a brief training period.

The auditory display should be easily heard in the OR alongside other sounds [27], being sufficiently distinguishable from but not interfering with existing sounds, such as those form suction devices, anesthesia equipment, or ICU sounds. The auditory display should not sound similar to an alarm, as the transmission of intensity levels using auditory display is not an alarm, and surgeons could become annoyed when presented with sounds that are perceived to be unnecessarily urgent [9]. Indeed, unnecessarily urgent sounds can become quickly fatiguing [28], and common auditory signals in clinical settings have been shown to convey an unintended, inappropriate level of urgency [29]. Although IEC 60601-1-8 is an international standard that addresses alarms in OR and categorizes these into low, medium, and high priorities, implementation of the standard is controversial [30-32]. Adherence to the standard is not compulsory although the medical alarm manufacturers are likely to comply [32]. The current work, however, is distinguished as an auditory display to be used actively by the surgeon, as opposed to an alarm, which merely notifies clinical staff at some point when a threshold has been reached. Thus, the auditory display for HHF probe measurement does not fall under the auspices of IEC 60601-1-8. Currently, there is no accepted standard for the design of non-alarm auditory displays in the OR, likely due to the paucity of investigations into the field of auditory display in interventions [16].

Although the preliminary results of this study are very promising, there are open questions regarding the implementation of such an auditory display system in the OR. First, although this specific application categorizes intensity values into four different levels (zero signal, low signal, high signal, and error), auditory display could also be enhanced to accommodate either fewer or more intensity levels. This number could dynamically change depending on the current clinical requirements or familiarity of the clinician with the auditory display. For

instances in which the clinical scenario is straightforward and requires only a binary decision support, the auditory display could, for example, be reduced to provide simply zero- and high-level tones. However, when the scenario is more complex, additional levels could be provided. The most relevant application for continuous sound is for fluorescence measurement during stereotactic biopsy [33,34]. In this application, it is desired to map the tumor, i.e., to provide information on the availability and intensity of the fluorescence along the stereotactic insertion path where the sites with the highest fluorescence signal are selected for biopsy. This could provide a stronger measure of confidence for these areas for which the clinician must be sure of eligibility for removal. A second option would be to encode the intensity using a hybrid approach that combines continuous mapping with the employed level-based tones. Thus, for instance, 1, 2, or 3 primary tones (such as those described in this work) could be augmented with secondary auditory parameter mapping such as the intensity of vibrato (slight frequency modulation) or tremolo (slight volume modulation). This could strike a balance between a quick classification of fluorescence intensities as well as provide clinicians with a more nuanced way of detecting critical areas at tumor boundaries. In any case, the clinical scenario and clinician familiarity with the auditory display should determine the appropriate method to be employed. Issues of certification must be addressed as well, as a binary classification would place a higher burden on the device manufacturer, whereas a continuous or multi-level approach would place a higher burden on the clinician to discriminate between tumor and healthy tissue.

In terms of auditory display design, the most frequently incorrectly perceived sound in the study (zero signal) could be redesigned to be more unique, thereby possibly increasing recognition rates by participants. Second, the latency of the fluorescence measurement might be accelerated so that the surgeon could receive intensity information even faster. Third, although the auditory display was specifically developed to produce tones that should not interfere with common existing OR sounds, the integration of the auditory display in the OR must be further customized to the specific environment in which it is to be played. This should minimize interference with sounds from other devices. Integration must also take playback mechanisms into account; for instance, a small loudspeaker placed locally near the surgeon could play back the tones rather than a speaker placed somewhere further away in the room.

4.3 Future directions

Future development should take the aforementioned factors into account and evaluate the auditory display in an environment that is even closer to that of the actual clinical scenario or in the OR. In addition, time, accuracy, and workload aspects should be compared to those cases in which an assistant relays intensity values to the surgeon to better discern the benefit of such an auditory display. Methods such as augmented reality and surgical microscope image injection will surely play a major role in the operating rooms of the future. However, this is often accompanied with negative impacts on clinician attention, for instance, significantly reducing foreign body recognition in endoscopic viewing supported by augmented reality [35].

The development of a hybrid display system could be beneficial, for instance, by transmitting the intensity of the quantified fluorescence with both auditory and visual means. Thus, future evaluations should determine in which cases each method of feedback is most suitable. To the authors knowledge, there have been no such comprehensive evaluations in the clinical field. An ideal subsequent solution could be a harmonization in a hybrid system that provides both immediate, relevant intensity information through auditory display and more complex details using visual display, either through a screen or using augmented or virtual reality concepts.

5 Conclusion

The experiments show that the employed auditory display is fairly intuitive, easy to learn, accurate, and fast in providing the user with the measurement of the current intensity or error signal for fluorescence-guided resection. Auditory display is a nascent field that could bring real benefit to surgeons using FGR, reducing reliance on a computer screen or surgical assistant and allowing focus to be retained on the surgical site. Future work should refine auditory methods and evaluate the concept in a more realistic clinical environment as well as investigate possible combinations with visual methods such as augmented and mixed reality.

Acknowledgments

The authors would like to thank Johan Richter, neurosurgeon at the Department of Neurosurgery in Linköping University and the participants for feedback on the sound system.

Funding

The study was supported by Swedish Childhood Cancer foundation (Grant No. MT 2013-0043), the Cancer network at Linköping University (LiU-cancer) and National Institutes of Health Grants P41 EB015902, P41 EB015898, R01EB014955, and U24CA180918.

References

- Stummer W, Pichlmeier U, Meinel T, Wiestler OD, Zanella F, Reulen H-J. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. Lancet Oncol. 2006; 7(5):392–401. DOI: 10.1016/ S1470-2045(06)70665-9 [PubMed: 16648043]
- Haj-Hosseini N, Richter J, Andersson-Engels S, Wårdell K. Optical touch pointer for fluorescence guided glioblastoma resection using 5-aminolevulinic acid. Lasers Surg and Med. 2010; 42(1):9–14. DOI: 10.1002/lsm.20868 [PubMed: 20077492]
- Haj-Hosseini N, Richter J, Hallbeck M, Wårdell K. Low dose 5-aminolevulinic acid: implications in spectroscopic measurements during brain tumor surgery. Photodiagn Photodyn Ther. 2015; 12(2): 209–214.
- Richter J, Haj-Hosseini N, Andersson-Engel S, Wårdell K. Fluorescence spectroscopy measurements in ultrasonic navigated resection of malignant brain tumors. Lasers Surg and Med. 2011; 43(1):8–14. DOI: 10.1002/lsm.21022 [PubMed: 21254137]
- Richter J, Haj-Hosseini N, Hallbeck M, Wårdell K. Combination of hand-held probe and microscopy for fluorescence guided surgery in the brain tumor marginal zone. Photodiagn and Photodyn Ther. 2017; 18:185–192. DOI: 10.1016/j.pdpdt.2017.01.188
- Utsuki S, Oka H, Miyajima Y, Shimizu S, Suzuki S, Fujii K. Auditory alert system for fluorescenceguided resection of gliomas; Technical Note. Neurol Med chir. 2008; 48(2):95–7. DOI: 10.2176/ nmc.48.95

- Sanderson P, Watson M, Russell W. Advanced patient monitoring displays: tools for continuous informing. Anesth Analg. 2005; 101:161168.doi: 10.1213/01.ane.0000154080.67496.ae
- Woerdeman P, Willems P, Noordmans H, van der Sprenkel J. Auditory feedback during frameless image-guided surgery in a phantom model and initial clinical experience. Neurosurgery. 2009; 110:257–262. DOI: 10.3171/2008.3.17431
- Willems P, Noordmans H, van Overbeeke J, Viergever M, Tulleken C, van der Sprenkel J. The impact of auditory feedback on neuronavigation. Acta Neurochirur. 2005; 147:167–173. DOI: 10.1007/s00701-004-0412-3
- Voormolen E, Woerdeman P, van Stralen M, Noordmans H, Viergever M, Regli L, van der Sprenkel J. Validation of exposure visualization and audible distance emission for navigated temporal bone drilling in phantoms. PLoS ONE. 2012; 7(7):e41262.doi: 10.1371/journal.pone. 0041262 [PubMed: 22848452]
- Cho B, Oka M, Matsumoto N, Ouchida R, Hong J, Hashizume M. Warning navigation system using real-time safe region monitoring for otologic surgery. Int J Comput Assist Radiol and Surg. 2013; 8(3):395–405. DOI: 10.1007/s11548-012-0797-z [PubMed: 23161169]
- Hansen C, Black D, Lange C, Rieber F, Lamadé W, Donati M, Oldhafer K, Hahn H. Auditory support for resection guidance in navigated liver surgery. Med Robot and Comput Assist Surg. 2013; 9(1):36.doi: 10.1002/rcs.1466
- Bork, F., Fuerst, B., Schneider, A., Pinto, F., Graumann, C., Navab, N. Auditory and visio-temporal distance coding for 3-dimensional perception in medical augmented reality. Proceedings of 2015 IEEE international symposium on mixed and augmented reality (ISMAR); 2015. p. 7-12.
- Black D, Hettig J, Luz M, Hansen C, Kikinis R, Hahn H. Auditory feedback to support imageguided medical needle placement. Int J Comput Assist Radiol and Surg. 2017; doi: 10.1007/ s11548-017-1537-1
- Strauß G, Schaller S, Zaminer B, Heininger S, Hofer M, Manzey D, Meixensberger J, Dietz S, Luth T. Klinische Erfahrungen mit einem Kollisionswarnsystem. HNO. 2010; 59:470479.doi: 10.1007/s00106-010-2237-0
- Black, D., Hansen, C., Nabavi, A., Kikinis, R., Hahn, H. A Survey of auditory display in imageguided interventions; Int J Comput Assist Radiol and Surg; 2017.
- Wright, M., Freed, A., Momeni, A. OpenSound Control: State of the Art 2003. Proceedings of 2003 international conference on new interfaces for musical expression (NIME); 2003. p. 153-159.
- Haj-Hosseini N, Kistler B, Wårdell K. Development and characterization of a brain tumor mimicking fluorescence phantom. Proc SPIE. 2014; 8945:894505-894505-894507. DOI: 10.1117/12.2039861
- Blattner M, Sumikawa D, Greenberg R. Earcons and icons: their structure and common design principles. Hum Comput Interact. 1989; 4(1):11–44. DOI: 10.1207/s15327051hci0401_1
- DuBois, T. Doctoral Dissertation. University of Southern California; Los Angeles: 1983. Christian Friedrich Daniel Schubart's Ideen Zu Einer Ästhetik Der Tonkunst: An Annotated Translation; p. 1-33.
- 21. Puckette, M. Second intercollege computer music concerts. 1996. Pure data: another integrated computer music environment; p. 37-41.
- 22. Haahr, M. Website. School of Computer Science and Statistics; Trinity College, Dublin: 2016. http://www.random.org/sequences/ [Accessed Dec 2016]
- Kim A, Khurana M, Moriyama Y, Wilson B. Quantification of in vivo fluorescence decoupled from the effects of tissue optical properties using fiber-optic spectroscopy measurements. J Biomed Opt. 2010; 15(6):067006.doi: 10.1117/1.3523616 [PubMed: 21198210]
- Aalders M, Sterenborg H, Stewart F, van der Vange N. Photodetection with 5-aminolevulinic acid induced protoporphyrin IX in the rat abdominal cavity: drug-dose-dependent fluorescence kinetics. Photochem and Photobiol. 2000; 72(4):521–525. DOI: 10.1562/0031-8655(2000)0720521pwaaip2.0.co2
- 25. Stummer W, Tonn J, Goetz C, Ullrich W, Stepp H, Bink A, Pietsch T, Pichlmeier U. 5-Aminolevulinic acid-derived tumor fluorescence: the diagnostic accuracy of visible fluorescence qualities as corroborated by spectrometry and histology and postoperative imaging. Neurosurgery. 2014; 74(3):310–320. DOI: 10.1227/NEU.00000000000267 [PubMed: 24335821]

- 26. Eljamel S, Petersen M, Valentine R, Buist R, Goodman C, Moseley H, Eljamel S. Comparison of intraoperative fluorescence and MRI image guided neuronavigation in malignant brain tumours, a prospective controlled study. Photodiagn and Photodyn Ther. 2013; 10(4):356–361. doi:http:// dx.doi.org/10.1016/j.pdpdt.2013.03.006.
- 27. Edworthy J, Hellier E. Alarms and human behaviour: implications for medical alarms. Br J of Anaesth. 2006; 97(1):12–17. DOI: 10.1093/bja/ael114 [PubMed: 16698858]
- 28. Parseihian G, Ystad S, Aramaki M, Kronland-Martinet R. The process of sonification design for guidance tasks. J Mob Med. 2015; 9(2)
- 29. Mondor T, Finley G. The perceived urgency of auditory warning alarms used in the hospital operating room is inappropriate. Can J of Anaesth. 2003; 50(3):221–228. DOI: 10.1007/bf03017788 [PubMed: 12620942]
- Sanderson P, Wee A, Lacherez P. Learnability and discriminability of melodic medical equipment alarms. Anaesthesia. 2006; 61(2):142–147. DOI: 10.1111/j.1365-2044.2005.04502.x [PubMed: 16430567]
- Cvach M. Monitor alarm fatigue: an integrative review. Biomed Instrum & Technol. 2012; 46(4): 268–277. DOI: 10.2345/0899-8205-46.4.268
- 32. Edworthy J. Medical audible alarms: a review. J Am Med Inform Assoc. 2013; 20(3):584–589. DOI: 10.1136/amiajnl-2012-001061 [PubMed: 23100127]
- Haj-Hosseini, N., Richter, J., Milos, P., Hallbeck, M., Wårdell, K. Optical guidance for stereotactic brain tumor procedures - preliminary clinical evaluation. Proceedings of Photonics West; San Francisco. 2017; 2017.
- Markwardt N, von Berg A, Fiedler S, Goets M, Haj-Hosseini N, Polzer C, Stepp H, Zelenkov P, Rhm A. Optical Spectroscopy for Stereotactic Biopsy of Brain Tumors. In proceedings SPIE. 2015; 9542(954208):1–8. DOI: 10.1117/12.915751
- 35. Dixon B, Daly M, Chan H, Vescan A, Witterick I, Irish J. Surgeons blinded by enhanced navigation: the effect of augmented reality on attention. Surg Endosc. 2013; 27:454–461. DOI: 10.1007/s00464-012-2457-3 [PubMed: 22833264]







Fig. 1.

Information communication scenario in the OR (a), showing a surgeon manually placing the HHF probe with the right hand on the surgical site while viewing a display in approximately 4 m away, and (b) examples of fluorescence spectra measured during surgery that are converted to zero, low, high, and error tones.

			Set 1			1	Set 2			!	Set 3		
Fiber-optic	2	Е	1	0	2	Е	0	1	0	2	Е	1	
probe	1	2	Е	0	Е	0	2	1	0	Е	1	2	
	1	0	Е	2	Е	0	1	2	0	1	Е	2	
	1	Е	2	0	1	Е	0	2	2	0	1	E	
	Е	2	0	1	2	0	Е	1	Е	1	0	2	
A A A A A A A A A A A A A A A A A A A	2	1	0	Е	0	1	2	Е	0	Е	2	1	
	0	2	1	Е	1	2	0	Е	1	0	2	Е	
	Е	2	1	0	2	1	Е	0	Е	1	2	0	
(a)							(b)						

Fig. 2.

a) Experimental setup and the phantoms showing HHF probe and tray of liquid phantoms and b) the phantom arrangement with three sets of 32 wells, where the content of the wells is indicated by 0 (zero signal), 1 (low signal), 2 (high signal), or E (error signal). The sequence followed by the participants is shown with arrows.



Fig. 3. Flowchart of signal generation in LabVIEW



Fig. 4.

Profile of the auditory display tones, where frequency of individual oscillators is shown on the y-axis, playback time from 0 to 1000 ms on the x-axis. Amplitude envelope is depicted using a gradient from white to dark gray. The zero signal tone employs sine oscillators, and the remaining tones employ triangle oscillators. a) Zero signal, b) low signal, c) high signal, d) error signal



Fig. 5.

Latency variations for each participant. The overall response accuracy for each participant is annotated as percentages in the upper row of the graph.



Fig. 6.

Median of latency versus accuracy for each participant. No correlation was found between the two parameters. The musical proficiency of the participants is notated beside each data point, where 0 is no musical training, 1 is amateur training, and 2 is professional training.

Table 1

Phantom composition;

Phantoms	AF	PpIX (g/l)	Sig. Levels
Zero	Yes	0	0–0.9
One	Yes	10	1.0–12.9
Two	Yes	30	13.0-50.0
Error	Yes*	0	5000

* Autofluorescence blocked by the addition of extra ink.

Confusion matrix for intuition test results showing played versus perceived functions of the auditory display tones

			Pla	iyed	
		Zero	One	owT	Error
	Zero	3	9	0	2
Deriver	One	5	2	0	2
rerceived	Two	0	0	10	0
	Error	2	2	0	9

Author Manuscript

Table 3

Confusion matrix for function response accuracy showing played versus perceived functions of the auditory display during laboratory evaluation measurement

Black et al.

			Pla	iyed	
		Zero	One	Two	Error
	Zero	247	2	0	0
Douton	One	7	258	1	1
navianiai	Two	0	3	267	1
	Error	1	0	1	171