

Point Estimation with Markers for Effective Mobile Auditory Graphs

<https://doi.org/10.3991/ijet.v15i08.12347>

Zico Pratama Putra ^(✉), Deni Setiawan
Queen Mary University of London, London, United Kingdom
z.putra@qmul.ac.uk

Bagus Priambodo
Universitas Mercu Buana, Jakarta, Indonesia

Mila Desi Anasanti
Imperial College London, London, United Kingdom

Abstract—While researchers have performed numerous studies to understand the human interpretation of visual graphs in reading, comprehending and interpreting displayed data; visually impaired (VI) students still face many challenges that prevent them from fully benefiting from these graphs in class. In this study, we conducted a test with 20 students to track the work described in studies in an expanded scenario. As we have tried to answer the question as to whether adding multi-reference mapping of sonification to auditory graphs could improve the of point estimation accuracy in non-visual condition. We also emphasize the efficiency of the performance of multi-reference graphs to make them as efficient as mapping using single pitch. Our proposed study improved performance of multi-reference task completion time by having fewer reference note. The results help provide empirical evidence that the multi-reference mode provides more accurate results than the single-pitch mode and confirms that adding contexts to auditory graphs could be used for better comprehension.

Keywords—Auditory graph, collaborative work, ubiquitous and mobile devices

1 Introduction

The implementation of auditory displays, which deals with the use of non-speech sound to display information, has been engaged in many complex work environments, ranging from computer applications, medical workstations, aircraft cockpits, and control centres of atomic reactors. A specific type of auditory displays is sonification, which is a technique used to typically mapping data sets to acoustic parameters to represent the data audibly. It usually represents information in non-speech audio [1]. It can help users to analyse the trend of data and its distribution by hearing the sound as a representation of the rendered acoustic data [2]. Data sonification benefits from

the fact that it can be perceived more broadly and clearly than speech that is precise and requires more focus [3].

Many studies have been proposed to improve science, technology, engineering and mathematics (STEM) education for VI students using sounds in learning [4]; visually impaired (VI) users still face many challenges that prevent them from fully benefiting from graphs. Consequently, it affects their understanding of data visualisation, and in turn reduces their role in collaborative tasks with their sighted peers in both educational and working environments. Although, the auditory graphs have been proposed for these VI users, most studies of auditory graphs have only been carried out on a non-portable device like personal computer (PC) which does not offer specific modality input like haptic interaction.

In this study, we are interested in exploring further support of non-visual point estimation tasks using another form of sonification by integrating multiple tones as references to represent a note as previously conducted by Metatla et al [5]. Their study showed that using multiple references could improve the accuracy of point estimation tasks in auditory graphs.

2 Background

2.1 Auditory graph implementation on mobile devices

Research related to the auditory display on a mobile device has been evolving in recent years due to the extensive use of mobile communication [6]–[9]. Researchers have exploited the use of accessibility features such as screen readers and voice commands for communication purposes such as to create, record, send and receive emails; use of maps for navigation; and modify a document [8], [10], [11]. An earlier study has developed a system on a tablet PC, called exPLoring graphs at UMB (PLUMB), which was designed to support people with visual impairments to understand graph by using auditory cues [12]. To date, the closest work to our study has been developed by researchers from Monash University, GraCALC, as an approach for implementing numerical and statistical graphics to VI [13]. The system presents a graphic from a mathematical function as a line graph which then displayed on a web-based service. Putra et.al. have designed mobile auditory graph (MAG) to support collaborative task for VI users to enable them to create and edit graph collaboratively [14].

2.2 Auditory graph design

The design of auditory graphs focuses on the question of mapping the dimensions of sound to the displayed data. The main mapping issue includes whether pitches should be increased or decreased in response to changes in the associated data. Auditory graphs can be considered a class of sonified displays which uses sound to display quantitative information. This means that any changes in quantitative data are mapped to changes in one or more dimensions of sound. Auditory graph as part of auditory display framework may solve the audio clutter arises from an attempt to listen to

many numeric values in speech. Imagine the difficulty of trying to remember 10 or more data values spoken out loud. Non-speech sounds facilitate the traceability of auditory graphs by simply following a continuous sound trend whose pitch changes according to the values in the data set.

As in the visual graphs, the auditory graph characteristics need to be set up properly so that the listener can understand the meaning of data. While, the properties of the visual graphs (i.e., spatial area, colour, trend, and size) are regularly changed, the audio properties in sonification, such as pitch, pan, rate, volume, and timbre may be modified. These properties describe some mappings to the sound attributes such as loudness (identifies with the sound's amplitude), pitch (a feature that relates to the frequency of sound) and timbre (a characteristic of a sound that identifies it from the various reference of a similar pitch and volume) [15].

Walker et.al. [16] explored these questions by comparing data-to-display mappings, polarities and scaling functions to correlate data values with the associated sound parameters for both sighted and VI people. They discovered that in some circumstances, VI listeners might prefer opposite polarities to sighted listener perceptions. They found that for a specific mapping, for example, mapping coin size to pitch, VI people tended to have an opposite mapping compared to sighted people.

Brown et al. [17] explored further the question of mapping sound and formulated guidelines for the design of auditory graphs based on research into the sonification of line graphs and guidelines for sonifying graphs with two or three series of data. Earlier research suggests that adding context cues such as checkmarks and labels to graphs offers advantages for non-visual interaction [18], [19]. Recognizing the point's position in space plays an integral role in reading and/or building graphically based representations [20].

Metatla [5] study developed a simple user interface for the users to predict a point by providing a vertical slider that can be moved along the Y-axis with the two modes: single point display and multi-reference display. The users can estimate the point when positioning it at a desired location on an axis [5].

In the single point display or pitch-only display, by mapping the pitch of a sine wave to the Y coordinate of the point with a positive polarity, users predict each point's position on an axis, i.e. increasing the pitch upwards for each point and decreasing the pitch downwards for each point. Using the same positive polarity as the pitch-only display, the position of a point in the multi-reference display can be predicted relative to an origin with multi-reference of tones. This can be done by determining both the pitch difference on that point in comparison with the subsequent points and the length of the sum of the successive notes separating it from the origin. Thus, a greater distance provides a longer sequence of tones. An ascending set of tones will be produced for points located below the origin, and a descending set of tones for those above the origin [5].

Metatla et.al found that there was a major disadvantage of using the multi-reference mode as it takes more time as compared to the use of usual pitch only graphs [5]. Based on their study, we would like to test whether using less multi-references will lead it to be more efficient as users can retain less information to remember the data, yet still improve the accuracy of point estimation.

Therefore, our study investigates whether employing multiple tones can assist point estimation tasks for perceiving a better perception and interpretation in auditory graphs, yet still as efficient as those using pitch only graphs. In order to confirm this assumption, 20 sighted participants took part in the study in March 2019.

In our MAG app., we developed an algorithm to only works for positive Y values with ascending from 0 up to a maximum value (YMax) for multi-reference. The idea is to play notes in multiples of 10ths of the maximum Y value leading up to the value of the point the user is trying to estimate (YEstimate). This approach is used in order to reduce the number of reference tones presented while still giving the user enough information to make a fairly accurate estimate of the point of interest. The reason for this is that researchers [5] found that if a lot of reference tones were played, people lost track of them and they became less useful.

3 Experiment

We developed an experiment to study the influence of sonification on the accurate estimation of point positions by adding reference markers. Our research centered on providing information that could help estimate the point position in relation to its proximity to an initial point.

3.1 Apparatus

The MAG interface has been developed based on a 9.7-inch screen Samsung Galaxy Tab S2 with the Android 7 operating system. We designed a mobile graph application with multimodal input by allowing gesture interaction to support the task of rendering the auditory graph (see Figure 1). Swipe interaction was implemented to help the users to locate the points on the X-Y coordinates.

Pitch-only mapping: We sonified the position of a point on an axis in the first design by plotting the pitch of piano and coin tone to the Y coordinate of the point according to a positive polarity. The tone's pitch changes in accordance with the point's movements on the axis; the value with the lowest frequency was mapped into coin sound and piano for higher frequency and was increased linearly to attain the maximum value of 100 in 1638 Hz frequencies in midi note G#6.

Multi-references mapping: In the second design, we applied the same pitch mapping used in the first prototype. But instead of hearing only one pitch, there are sequential reference tones with different pitches that correspond to all points before the current position and the original reference. This version only works for positive Y values. We think of the Y values going from 0 up to a maximum value (YMax). The idea then is to play notes leading up to the value of the point the user is trying to estimate (YEstimate). Whether $YEstimate > YMax * 0.5$, we only start playing them from the point $YMax * 0.5$, so the user will never have more than 4 notes played before the value of YEstimate is played.

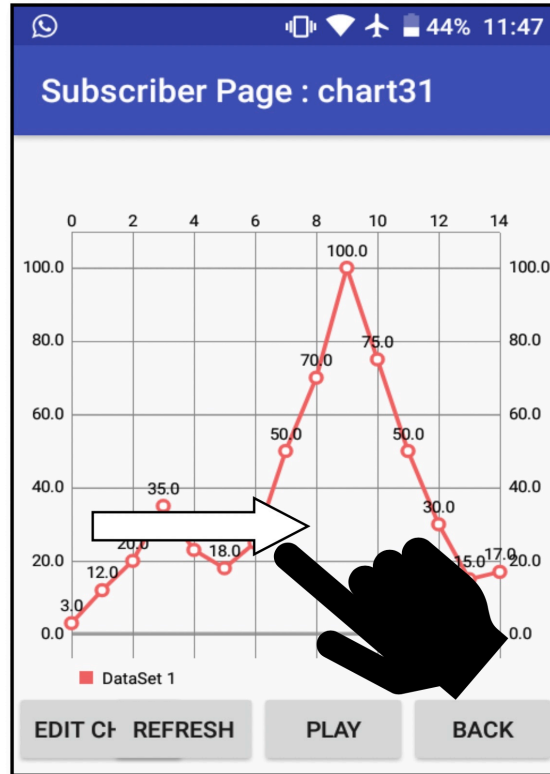


Fig. 1. Illustration of user’s hand interacting with the MAG interface by tapping and/or swiping from the left part of the graph to the right.

In order to generate a sequence of tones, our reference tones had a duration of 200 milliseconds and were superimposed with a delay of 200 milliseconds. The point's position can be estimated in relation to an origin by comparing the difference in pitch at that point with the next points and the total length of the successive notes separating it from the origin. A longer distance yields a longer succession of tones. It will be important to use a different instrument when playing fractions of $Y_{Estimate} > Y_{Max} * 0.5$ compared to the instrument used to play $Y_{Estimate} < Y_{Max} * 0.5$, as we clearly want to be able to distinguish between the two.

In our MAG app., whether $Y_{Estimate}$ is higher than half of Y_{Max} ($>0.5*Y_{Max}$) or below half of Y_{Max} ($<0.5*Y_{Max}$), the system only begins playing the reference in ascending order. For the point above half of Y_{Max} , the system plays the reference from the point above half of Y_{Max} with a piano sound, while for the point below half of Y_{Max} , the reference from the point above zero is played with coin sound. Thus, the user will never have more than 5 notes played including the value of $Y_{Estimate}$. For example, when the user reaches number 86 with Y_{Max} number is 100, he/she hears the sequence of tones composed of all the pitches of points 60, 70, 80 and 86 as the

YEstimate. When the user reaches number 46, he/she hears the sequence of tones composed of all the pitches of points 10, 20, 30, 40 and 46 as the YEstimate.

The sounds were generated by mapping the Y-coordinates of graphs into the pitches. In this study, the Y-axes were ranged from 0 to 100 which one value as one increment as shown in Figure 1.

3.2 Experimental design

Participants: A total of 20 sighted participants volunteered to take part in study 1 (16 men and 4 women) between 18 and 39 years old, and another 20 sighted participants in study 2 (11 men and 9 women) between 18 and 39 years old. They were a mixture of university staff (both academic and non-academic), undergraduate and postgraduate students from Queen Mary University of London. They were randomly assigned to two groups of ten in a within-subject experimental design.

Procedure: Participants were given an overview of the experiment on arrival as the aim of this study. They presented with the explanation related to how long the study will be performed and the instruction of the interface on how to assess the auditory graphs in our MAG app.

After that, they were asked to complete the first questionnaire about their demographic details and musical education (in relation to the musical instrument). Fourteen participants assessed their musical education as “playing no musical instrument” or beginner, the rest of them stated that they are playing at least one musical instrument. The participants had no experience with non-visual interaction.

The participants were then randomly assigned to one of the two groups. Each of them was asked to estimate two different graphs (graph A and graph B) with one of the two modes (pitch-only or multi-reference mode). They were paired to use different mode for each of the graphs so that it will diminish the learning effect using two modes with the same graph. For example, user 1 was asked to estimate the graph A with the pitch-only mode and the graph B with the multi-reference mode, while user 2 was asked to use the opposite mode using the multi-reference mode for the graph A and the pitch-only mode for the graph B.

Participants were trained on the respective displays and could spend as much time as they want to familiarize themselves with the interfaces before starting the experiments. Specifically, the participants were familiarized with the various sonification mappings used and asked to spend sufficient time until they felt comfortable with both the mappings used.

The training usually lasted 10 minutes per each sample graph with pitch-only or with multiple-reference markers. The training graphs was in a visual state, which was first displayed as a linear graph with notes from 10 to 100 in multiples of 10, i.e. 10, 20, 30, and so forth. The participants were then presented in the second graph in the form of randomly arranged notes, however, still have a maximum value of 100. The purpose of this second graph was not only to reinforce their memories of the range of pitch sounds, but also to familiarize them to the display of other values between tens, such as 18, 23, 75, etc. The participants were asked to lay their fingers on the left-hand side of the MAG app interface that showed the graph, then were asked to slide

their finger until the end of the right-hand side. When their finger touched a certain point, they heard a tone of the Y-value in accordance to the real value they had seen on the graph.

After the training, the testing task usually lasted from 4 to 6 minutes per condition in the non-visual state. In this setting, participants could see the graph area, however the line graph is hidden. The users had to rely on the sonifications to estimate the position of the points as they swiped their finger to the target positions. They were asked to perform 10-point estimation trials per mode, i.e. either pitch-only or multi-reference mode. Initially, participants began to tap the MAG app from the left part of the graph for the first trial and swiped slightly to the right direction for the next trial and so forth. They were asked to state their estimated numbers every time they hear a pitch to be noted by the observer. They were not given any feedback about the real Y-value after the test.

3.3 Research question

The main research questions of the experiment were:

1. Will users produce lower point estimation errors when using multi-reference sonification mapping relative to pitch-only sonification mapping?
2. Will it still be slower performing point estimation tasks when using the multi-references sonification mapping with fewer references no more than five tones as compared to the pitch-only mapping?

4 Results

4.1 Study 1

Evaluation of point estimation error with one sample audio: To evaluate how well users can perform point estimation task, we observed the RMSE between the estimated (predicted) values to the true values, apart from the different categories of graphs and different presentation modes. This is an exploratory study, with the intention of examining how well sighted participants perform point estimation tasks using the first version of the MAG app prototype.

The results were then calculated across all subjects by calculating the RMSE between the estimate values with the true values; one for the pitch only and the other for the multi-reference mode. This separation of using 2 modes was implemented as we were interested to know whether there is a relationship between the performance of point estimation tasks and the mode used to perform the tasks. To confirm whether the difference between the two modes is statistically significant, we performed a student t-test comparing the means of the RMSE obtained between the two modes. Our null hypothesis is that the RMSE mean of the pitch-only mode is equal to the mean of the multi-reference mode. A one-tailed test was used to test if the RMSE mean of pitch-only mode is significantly greater than those of the multi-reference mode.

The t-test resulting in an insignificant different of the RMSE means which implies that the two modes are not significantly different ($t = -1.59$, $p = 0.121$).

The experiment was failed to prove that the multi-reference mode provides more accurate results than the pitch-only mode.

Completion time: The completion time of point estimation tasks was calculated across all subjects; one for the pitch only and the other for the multi-reference mode. The t-test resulting in a significant difference in the completion time means which implies that the two modes are significantly different in regard to the time completion used ($t = 2.6719$, $p = 0.01134$). In term of completion time, the experiment was also failed to prove that the multi-reference mode performance could have same level to the pitch-only mode. Based in this result, we consider implementing two sample audio sources, having coin for point below half of YMax and the rest with piano sound as describe previously, as follows:

4.2 Study 2

Evaluation of point estimation error with two sample audios: The results of point estimation tasks using piano and coin sample audio calculated across all subjects by calculating the RMSE between the estimated values with the true values; one for the pitch only and the other for the multi-reference mode. This separation of using 2 modes was implemented as we were interested to know whether there is a relationship between the performance of point estimation tasks and the mode used to perform the tasks. After calculating the RMSE, the values were plotted into four boxplots to visualise the distribution of the error for each method and each type of graph.

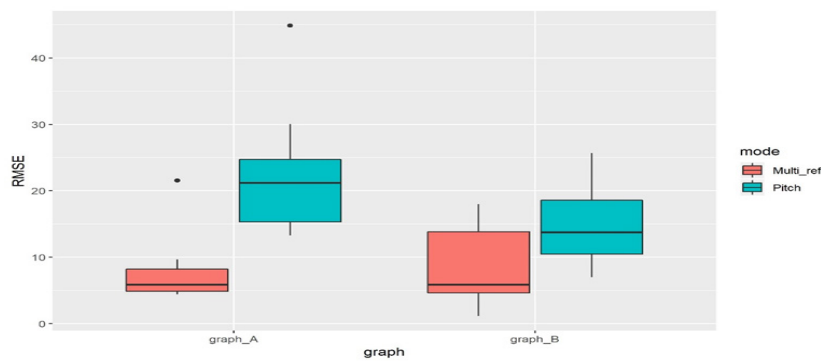


Fig. 2. Comparison of four boxplots, representing the distributions of RMSE obtained from the multi-reference mode and the pitch only in graph A and graph B as displayed on the X-axis. The Y-axis shows the RMSE values between each error from 0 to 30. The legend denotes the task in the multi-reference and the pitch only mode.

We then combined the RMSE results from graph A and graph B as both graphs were assigned with random values and patterns, thus can be treated as if they are in one graph. The boxplot of these combined graphs consistently shows that using multi-

reference mapping improved the performance represented by its lower RMSE values as shown in Figure 2.

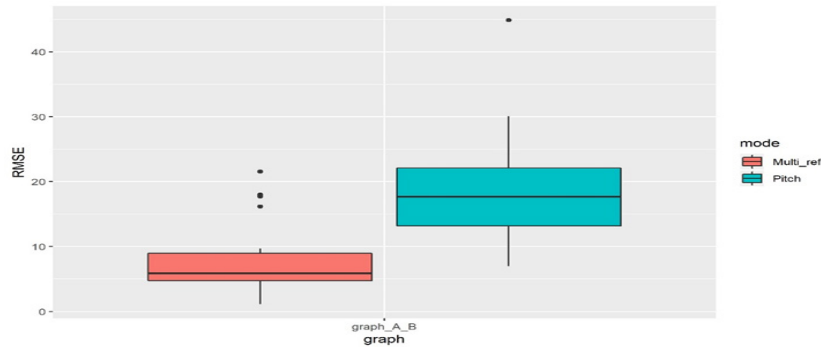


Fig. 3. Comparison of the boxplot of RMSE values calculated from graph A and graph B combined using the multi-reference and the pitch only mode.

To confirm whether the difference between the two modes is statistically significant, we performed a student t-test comparing the means of the RMSE obtained between the two modes. Our null hypothesis is that the RMSE mean of the pitch-only mode is equal to the mean of the multi-reference mode. A one-tailed test was used to test if the RMSE mean of pitch-only mode is significantly greater than those of the multi-reference mode.

The t-test resulting in an insignificant difference of the RMSE means which implies that the two modes are not significantly different ($t = -4.59$, $p = 0.00006.172$). The aim was achieved through this study; the experiment's results generally show that the multi-reference mode provides more accurate results than the pitch-only mode.

Completion time: The completion time of point estimation tasks was calculated across all subjects; one for the pitch only and the other for the multi-reference mode. After calculating the completion time, the values were plotted into four box plots to visualise the distribution of the completion time for each method and each graph. The time is presented on Y-axis in milliseconds.

As seen from the boxplots in Figure 4, in general, the time used to complete the tasks using multi-reference mode shows less performances with those using pitch only mode in term of their median and quantiles with the distribution may slightly different. In graph A and B, the multi-reference mode has a wider distribution than the pitch-only mode.

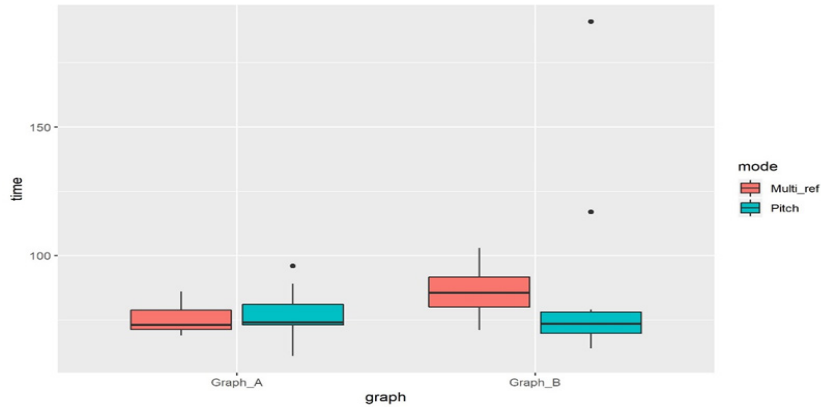


Fig. 4. Comparison of four boxplots representing the completion time from the multi-reference and the pitch only mode in graph A and graph B as displayed on the X-axis. The Y-axis shows the time in milliseconds (ms) from 0 to 200 ms. The legend denotes the task in the multi-reference (orange colour) and the pitch only (teal colour) mode.

Our null hypothesis is that the completion time of the pitch-only mode is equal to the mean of the multi-reference mode. A one-tailed test was used to test if the completion time mean of the pitch-only mode is significantly greater or smaller than those of the multi-reference mode.

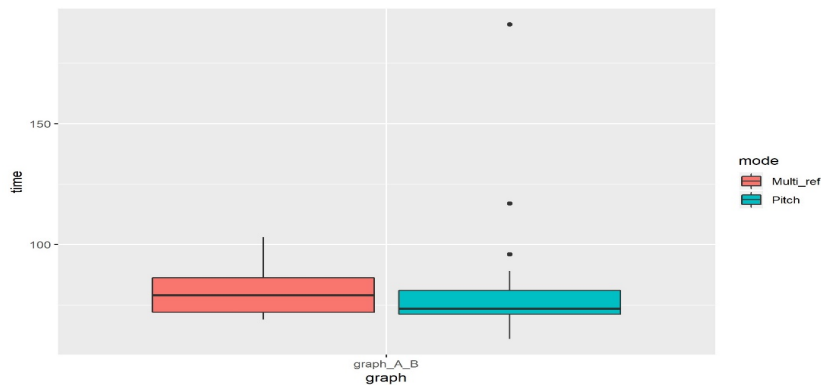


Fig. 5. Two boxplots depicting the completion time (in milliseconds) obtained to complete point estimation tasks using the multi-reference and the pitch only mode from the combined graph A and graph B.

When the completion time results from graph A and graph B were combined, the boxplot on Figure 5 consistently shows that completion time is not significantly different ($p > 0.05$), but the multi-reference has a wider distribution.

The t-test did not result in a significant difference of the completion time means which implies that the two modes are not significantly different in regards of the time completion used ($t = -0.299, p = 0.76$).

5 Discussion

5.1 Analysis of the accuracy to estimate point-estimation tasks

The exploratory study 1 was failed to prove that the multi-reference mode provides more accurate results than the pitch-only mode ($t = -1.59$, $p = 0.121$). The participants seemed to have difficulty differentiating markers before and after 50% of YMax while using only one sound sample (piano sound). They need additional references to separate the two parts of this pitch, so study 2 is needed under the same conditions as study 1, but include additional sample sound (coin sound) and split them for any number before and after 50% of YMax. In general, the results showed that the multi-reference mode generated more accurate results when the participants were asked to estimate for graph A and B. The t-test resulting a significant difference of the RMSE means which implies that the two modes are significantly different ($M_{\text{pitch-only}} = 18$, $M_{\text{multi-reference}} = 6$) ($t = -4.59$, $p = 6.172 \times 10^{-5}$). The first research question of this study has been answered for this population that the users produce higher point estimation errors when using the pitch-only sonification mapping compared to the multi-references sonification mappings.

5.2 Analysis of the completion time to finish the tasks

Concerning the duration of the trial, the users' opinions were divided between those who considered the pitch only mode to be faster and those who against it. In general, most participants considered the multi-reference was a faster for point estimation trial rather than pitch only. Their feedbacks were opposing the literature that it requires longer time for the multi-reference task completion. However, the pitch only mode could be the most demanding attention because they have no context information to help the point estimation process. Tough, most participants believed that the trial estimation time to complete the tasks was shorter in the multi-reference mode.

Most participants found it was really difficult to follow the note-by-note presentation in the pitch only mode. Although a few users felt that the completion time used to conduct the test on multi-reference graphs were longer than the one used pitch-only graphs, however, on average there was no significant difference in the duration of the test time between the two modes as shown in Figure 5. The conclusion seems to be different from Metatlas' [5] conclusion as they found that there was a compromise between speed and accuracy for multi-reference sonification. This tradeoff is actually could be anticipated by limiting the number of references. We successfully proved this claim in our experiment by creating the number of the references to a maximum of 5, separating the sound reference to under 50% and above 50% of YMax, and shortening the sound delay between the references. With this, participants will have a shorter time to maps the tones to the value and also the application will have a shorter time to present the graphs in tones as compared with the point presented in the work of Metatla et al. Therefore, research question 2 of this study has been answered, there was no significant difference between the task completion times between pitch-only

and multi-reference modes. In conclusion, most of the participants responded that multi-references were faster in the estimation trial, although the statistical test calculating the completion time between the two modes confirmed that the difference was not significant ($t = -0.299$, $p = 0.76$).

6 Conclusion

In general, the results of the experiment show that the multi-reference mode generated more accurate results compare to the pitch-only mode. The evaluation confirms that adding context to auditory graphs such as tick marks could enhances the perception of auditory graphs.

Furthermore, while a few participants considered that the point-only is faster but less accurate, most of the participants responded that the multi-reference mappings were faster in the point estimation trial. About 80% of them also considered that the multi-reference mode was also easier for estimation as the single note may require each user to remember the wider band of pitch sounds.

Compare to Metatla's work, we showed the improved performance on point estimation tasks containing fewer reference tones which results in less reduction in speed of performance of the point estimation tasks. We also successfully managed the approach to be scalable, in the sense that there will never be more than a fixed number of tones played no matter what the numbers are on the Y-axis.

Although the outcomes of this study have been successful, several areas require further investigation since there are several differences from Metatla's study: the fact Oussama used negative numbers and the polarity of the pitch change. Therefore, we plan to conduct further study more like Oussama condition but preserving the key aspect of our approach. The other condition could be similar to Metatla's, but not use negative notes, and it should use the same instruments and graphs as we did for the study 2. Since we have split the sample sound to coin and piano sound in the latest study, the coin sound should be replaced with piano notes throughout the whole display in future study.

7 Acknowledgement

This report was funded by the Indonesia Endowment Fund for Education, LPDP.

8 References

- [1] B. N. Walker and M. A. Nees, "Theory of Sonification," *Princ. Sonification An Introd. to Audit. Disp. Sonification*, pp. 1–32, 2011.
- [2] M. A. Garcia-Ruiz, A. Edwards, R. Aquino-Santos, M. Vargas Martin, and S. A. El Seoud, "Auditory display as a tool for teaching network intrusion detection," *Int. J. Emerg. Technol. Learn.*, 2008.

- [3] S. A. Brewster, “NonSpeech Auditory Output,” in *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*, Hillsdale: L. Erlbaum Associates Inc., 2002, pp. 220–239. <https://doi.org/10.1201/b11963-ch-10>
- [4] A. S. Drigas and R. E. Ioannidou, “Special education and ICTs,” *Int. J. Emerg. Technol. Learn.*, vol. 8, no. 2, pp. 41–47, 2013. <https://doi.org/10.3991/ijet.v8i2.2514>
- [5] O. Metatla, N. Bryan-kinns, T. Stockman, and F. Martin, “Sonification of Reference Markers for Auditory Graphs: Effects on Non-Visual Point Estimation Tasks,” no. April, 2016. <https://doi.org/10.7287/peerj.preprints.1376>
- [6] G. G. Lin and J. G. Scott, “A mobile phone system to find crosswalks for visually impaired pedestrians,” vol. 100, no. 2, pp. 130–134, 2012.
- [7] H. Shin, J. M. Lim, J. U. Lee, and K. U. Kyung, “Background display for visually impaired people in mobile touch devices,” *Dig. Tech. Pap. - IEEE Int. Conf. Consum. Electron.*, pp. 464–465, 2013. <https://doi.org/10.1109/icce.2013.6486977>
- [8] J. Su, A. Rosenzweig, A. Goel, E. de Lara, and K. N. Truong, “Timbremap: : Enabling the Visually-Impaired to Use Maps on Touch-Enabled Devices,” *Proc. 12th Int. Conf. Hum. Comput. Interact. with Mob. devices Serv. - MobileHCI '10*, pp. 17–26, 2010. <https://doi.org/10.1145/1851600.1851606>
- [9] J. Kamaghe, E. Luhanga, and M. Kisangiri, “The Challenges of Adopting M-Learning Assistive Technologies for Visually Impaired Learners in Higher Learning Institution in Tanzania,” *Int. J. Emerg. Technol. Learn.*, vol. 15, no. 01, pp. 140–151, 2020. <https://doi.org/10.3991/ijet.v15i01.11453>
- [10] P. Strumillo, “Electronic interfaces aiding the visually impaired in environmental access, mobility and navigation,” *3rd Int. Conf. Hum. Syst. Interact. HSI'2010 - Conf. Proc.*, pp. 17–24, 2010. <https://doi.org/10.1109/hsi.2010.5514595>
- [11] A. Hartanto and N. Ismawati, “Pembuatan Aplikasi Buta warna untuk penderita buta warna,” *J. Ilm. FIFO*, vol. 10, no. 2, p. 42, 2019. <https://doi.org/10.22441/fifo.2018.v10i2.005>
- [12] R. F. Cohen, R. Yu, A. Meacham, and J. Skaff, “PLUMB: Displaying Graphs to the Blind Using an Active Auditory Interface,” *Assets'05*, pp. 182–183, 2005. <https://doi.org/10.1145/1090785.1090820>
- [13] C. Goncu and K. Marriott, “GraCALC: An Accessible Graphing Calculator,” *Proc. 17th Int. ACM SIGACCESS Conf. Comput. Access.*, pp. 311–312, 2015. <https://doi.org/10.1145/2700648.2811353>
- [14] Z. P. Putra, D. Setiawan, B. Priambodo, M. Desi Anasanti, A. H. Hazidar, and E. Naf'an, “Designing a Multimodal Graph System to Support Non-Visual Interpretation of Graphical Information,” *J. Phys. Conf. Ser.*, vol. 1339, p. 012059, Dec. 2019. <https://doi.org/10.1088/1742-6596/1339/1/012059>
- [15] K. V. Nesbitt and S. Barrass, “Finding trading patterns in stock market data,” *IEEE Comput. Graph. Appl.*, vol. 24, no. 5, pp. 45–55, 2004. <https://doi.org/10.1109/mcg.2004.28>
- [16] B. N. Walker, G. Kramer, and D. M. Lane, “Psychophysical scaling of sonification mappings,” *Proc. ICAD 2000*, pp. 99–104, 2000.
- [17] L. M. Brown, S. A. Brewster, R. Ramloll, M. Burton, and B. Riedel, “Design Guidelines for Audio Presentation of Graphs and Tables,” *Psychology*, pp. 1–5, 2003.
- [18] M. A. Nees and B. N. Walker, “Encoding And Representation Of Information In Auditory Graphs: Descriptive Reports Of Listener Strategies For Understanding Data,” *Proc. 14th Int. Conf. Audit. Disp.*, pp. 8–13, 2008.
- [19] D. R. Smith and B. N. Walker, “Tick-Marks, Axes, And Labels : The Effects Of Adding Context To Auditory Graphs,” *Proc. 2002 Int. Conf. Audit. Disp.*, pp. 1–6, 2002.

- [20] E. R. Tufte, “The visual display of quantitative information,” IEEE Power Engineering Review. 1988.

9 Authors

Zico Pratama Putra and **Deni Setiawan** both work for School of Electronic Engineering and Computer Science, Queen Mary University of London, London, in United Kingdom

Bagus Priambodo works for Faculty of Computer Science, Universitas Mercu Buana, Jakarta, in Indonesia. Email: bagus.priambodo@mercubuana.ac.id

Mila Desi Anasanti works for Section of Genomics of Common Disease, Department of Medicine, Imperial College London, London, in United Kingdom

Article submitted 2019-11-13. Resubmitted 2020-01-11. Final acceptance 2020-01-13. Final version published as submitted by the authors.