Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 34 (2014) 360 – 367

Procedia
Computer Science

The 11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC-2014)

Touching the data: exploring data sonification on mobile touchscreen devices

Denis Nikitenko[a], Daniel Gillis[a]

[a]School of Computer Science, University of Guelph, Guelph N1G2T4, Canada

Abstract

This paper suggests using mobile touchscreen devices to assist students with vision loss in working with data. It presents an integrated approach that combines current sonification methods with interactive multi-touch gesture-based exploration of data, designed to aid students in mental visualization and comprehension of data and function plots. This approach aims to help students with vision loss study independently of support centres, collaborate with their peers, and participate in group studies. Initial user study evaluating this approach and demonstrating its feasibility is presented; further research pathways are also discussed.

Selection and peer-review under responsibility of Conference Program Chairs
Keywords: Sonification; auditory displays; multi-touch interfaces; assistive technologies; mobile applications

1. Introduction

Vision loss poses significant barriers to education. According to CNIB [1], only 45 percent of Canadians with vision loss graduate high school. This is in part due to learning, social, and physical barriers that these students face. One of these educational problems is the difficulty of presenting visual information to blind learners, since visual information representation plays an important role in mathematics and science education.

Currently, there are few options available for students with vision loss who study subjects that involve information visualization. Students can get help through their educational institutions and work with note takers. Relying on a note taker, however, limits students' ability for independent study and social integration, which forms a social barrier between them and their sighted classmates. There are also tactile tools available for students with visual impairments, such as braille printers. However, these tools are non-interactive. The user cannot, for example, zoom in on a graph to examine a specific region. In addition, such tools are typically only available at support centres provided by educational institutions. They are not accessible to students working from home or in a group study setting.

* Corresponding author. Tel.: +1-519-400-4473; fax: +1-519-837-0323.
E-mail address: dnikiten@uoguelph.ca

Selection and peer-review under responsibility of Conference Program Chairs
doi:10.1016/j.procs.2014.07.038

This paper investigates the use of mobile touchscreen devices for presenting data plots as sound and providing students and professionals with vision loss an alternative tool for mental visualization and comprehension of data. It relies on concepts developed in the field of sonification, broadly defined as the use of sound to communicate information[2].

The human auditory system is sensitive to multiple aspects of sound, such as loudness - variations in the intensity of the sound, pitch - the frequency of the sound, and timbre - number, relative level, and arrangement of spectral components in a sound. It can also locate of the sound source in three-dimensional space around the listener. Numerous studies have suggested that these characteristics can be used to transmit numerical information[3,4,5,6,7].

Humans associate pitch of a sound with the perceived vertical location of the sound source[3]. This naturally suggests an approach that maps pitch to numerical data values. This use of sonification to replace graphs is often called an audio graph. Use of audio graphs to help people with vision loss was suggested by Mansur et al.[3], who focused on sonifying linear plots of univariate functions by continuously varying pitch to represent motion along the x axis. This approach has been expanded by further studies, which suggested combining speech and non-speech sounds for teaching various mathematical concepts, including function plots[4]. More recent approaches[6,8] suggested incorporating interactive methodologies into auditory graph displays and allowing users to explore the data.

Multiple methods have been proposed for adapting auditory displays to assistive technology. One broad category includes the use of auditory graphs to represent mathematical functions and includes methods for important characteristics of a function, such as its derivative and series expansion[4,6]. Other approaches suggest representing scatterplots using sound[9] or a combination of audio and haptic methods[7].

Current research suggests that methods involving a combination of audio playback and user interaction have an advantage over those relying solely on audio. Such methods typically involve the use of force feedback devices[10] or specialized experimental wireless devices, such as the Tangible Active Objects used by Riedenklau et al.[7]. The disadvantage of such devices is that their lack of portability and potentially high cost restricts their deployment to specific locations, such as student support centres or students' homes.

This project, therefore, aims to develop a more portable methodology, as well as expand the ways in which users can interact with data using gestures. Its goal is to allow students with vision loss to participate in classroom and group study activities centred around visual data displays, and to facilitate wider access to sonification display devices. It combines audification displays with physical gestures detectable by commercially available mobile devices.

Touch screen devices have become common and there is a growing body of research on use of multi-touch gestures on such devices by visually impaired and blind users[11,12]. Touch screen devices are widely commercially available, portable, require no additional hardware, and could easily be deployed by students with vision loss in any setting, including the classroom. They can be used as audification displays by combining sound displays with multi-touch exploration of the displayed data and mathematical functions.

2. Proposed approach

The proposed approach involves two extensions to current audification methods. First, it suggests the use of 3D sound for audification, due to potential improvements to spatial localization of audified data offered by 3D sound engines. Second, it integrates multi-touch display functionality to allow users with vision loss enhanced interaction with data. Both of these extensions are described below.

The framework presented here is designed to have three modes (Figure 1), all of which rely on the same sonification technology, and all of which allow user interaction with audified data:

- Sonification of continuous mathematical functions
- Sonification of histograms
- Sonification of scatterplots

Since this project is aimed at helping students with vision loss work with data in group study and classroom settings, it is assumed that users will connect a pair of headphones to their mobile device. A pair of headphones only allows for a stereo mix and left-right channel separation. It does not typically allow for in-front/behind and above/below spatial localization of sound sources.

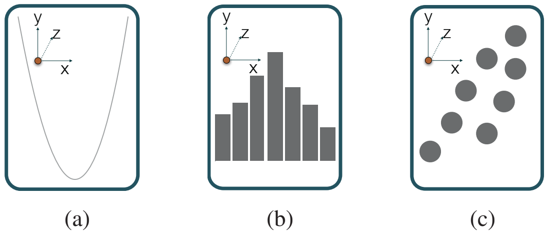


Fig. 1: Three sonification modes: function plot (a); histogram (b); scatterplot (c).

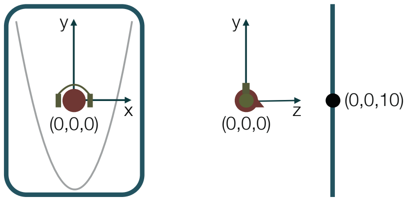


Fig. 2: 3D sound layout: the user is at the origin of the world space, facing the plane containing sonified data

The proposed sonification approach incorporates 3D sound to help overcome this limitation and aid spatial localization of sonified data. Spatial localization of audified data along the x-axis is typically done through stereo panning. Lower x values are panned left, and larger x values are panned right. A 3D sound representation allows this as well. A typical 3D sound environment consists of multiple sound sources and a single sound listener. Locations of sources relative to the listener result in different stereo representations of the sounds associated with the sources. As a result, a 3D sound engine naturally handles stereo panning of data.

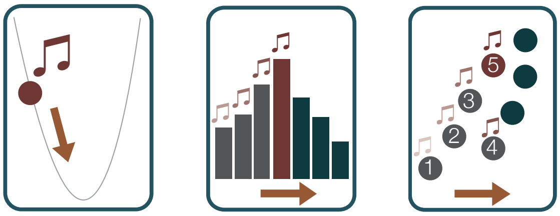
One potential advantage of 3D sound engines over simple stereo panning of data is that several currently available 3D sound engines provide support for head-related transfer functions (HRTFs). An HRTF characterizes how a human ear receives a sound from a point in space. HRTFs are used in pairs to simulate binaural sound, which appears to come from a specific point in space. Use of a 3D engine with HRTFs allows for enhanced vertical sound localization. This, combined with framework suggested below, would help users build a better model of the behaviour of a function or overall "shape" of an audified dataset.

Data sonification with 3D sound necessitates representation of data samples in several different spaces. Discrete data, and discrete samples of continuous functions, are stored in 2D data space. They are converted to two different spaces for presentation to the user: 3D audio space for audification, and 2D screen space for visual display on the touch screen. Conversion to 2D screen space is device- and API-dependent and is therefore omitted here for brevity. Data is converted into 3D audio space as follows: for every data point (x_{data}, y_{data}), a three-dimensional data point is defined as $\mathbf{p} = (x_{data} - x_{mid}, y_{data} - y_{mid}, 10)$, where $x_{mid} = (x_{max} - x_{min})/2$ and $y_{mid} = (y_{max} - y_{min})/2$, while x_{min}, x_{max}, y_{min}, and y_{max} represent the minima and maxima of the values being displayed.

The listener is placed into the origin of the 3D space (Figure 2). The listener's orientation is aligned with the world coordinate axes. The listener's up direction is aligned with the y-axis, forward - with the z-axis, and right - with the x-axis. Audified data presented on a vertical plane, parallel to the xy plane, in front of the listener (Figure 2). The plane containing 2D data is 10 units away from the listener and the displayed data is always centered relative to the listener position. The user interacts with the data using multi-touch gestures, as described below.

2.1. Sonification of univariate functions and bar plots

Univariate functions are sampled on a specific interval, creating an array of points in 2D data space that is sorted from smallest x coordinate to the largest. These points are converted into 3D audio space. A 3D sound source playing a continuous sound is initially positioned at the leftmost sample point. The source is translated from one 3D sample location to another, moving left to right along the x-axis, from the smallest x coordinate to largest (Figure



(a) (b) (c)

Fig. 3: Initial audification pass: function plot (a); histogram (b); scatterplot (c). The current audified value is represented in red. In a scatterplot, for points with equal x values, such as points 4 and 5, the point with a lower y (point 4) is audified before a pointed with a larger y (point 5).

3a). Audification loop pauses at each sample location for a short period of time and the total running time of the audification pass is set to three to five seconds, depending on the complexity of the function - in particular, presence of vertical asymptotes and discontinuities.

A sound sample plays continuously during the audification pass. Sound pitch is altered to represent the y position of the sample point being audified. A two-tone sound is used, with the base frequency of 400Hz and a 200Hz subcomponent. Pitch is scaled at each sample location by the y coordinate of the sample point: $scale = 5 * (y - y_{min})/(y_{max} - y_{min}) + 1$. If y is negative, the sound is processed though a low-pass filter to provide an audible difference between positive and negative values, as suggested by Grond et al.[6]. A distinct sound is also used to represent intersections of the graph with the x-axis. For piecewise continuous functions, continuous sounds separated by periods of silence.

Histograms are sonified using a similar approach. The only difference is that continuous sound samples are not used. Instead, a distinct click is played for each data point or bin (Figure 3b). The heigh of each bar - i.e. the number of data points in a bin - is used to scale the pitch of the click using the scaling factor calculated above.

Three-dimensional sound engines are typically developed for gaming applications, where physically accurate modelling of sound is not always necessary. As a result, 3D sound software provides several models for attenuating volume of a sound based on the distance from the listener; these are referred to as “rolloff” models. For the purposes of data sonification, attenuation of sound due to the distance from the listener is undesirable and 3D sound rolloff is therefore disabled. All data points are equally loud, regardless of how close they are to the listener.

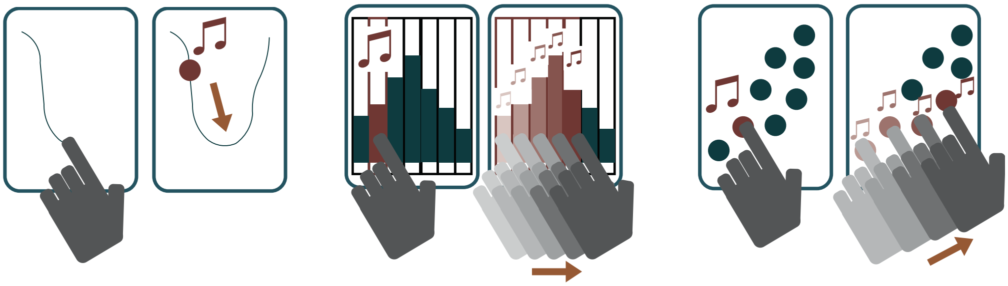
2.2. Sonification of scatterplots

For sonification of scatterplot, data is sorted in increasing order along the x-axis. Data points with the same x coordinate are sorted and audified in increasing order along the y-axis (Figure 3c). For example, in Figure 3c, points are audified for the user in the sequence 1, 2, 3, 4, 5, with point 4 being audified before point 5. Similarly, a scatterplot containing data points $(3,4),(4,3),(1,3),(3,2),(2,4)$ will be audified in the following order: $(1,3),(2,4),(3,2),(3,4),(4,3)$. Audification of individual data points is done using discrete clicks. Panning and scaling of the sound are implemented as described above. The listener hears a stream of clicks of various frequency, somewhat akin to a tune played on a xylophone, panning from left to right, and either increasing or decreasing in pitch, which is meant to create a mental picture corresponding to the overall shape of the scatterplot.

2.3. Interactive exploration of sonified data

Sonification is different from visual display in that it is intrinsically transient. As a result, data may need to be sonically displayed several times to allow users to form a better mental model of the data. Transience of sonification can also be mitigated by physically exploring the data, as suggested, for example, by Riedenklau et al.[7].

Current commercially available touchscreen technology does not yet allow for tactile feedback to touches. However, the combination of multi-touch screens and sound feedback afforded by modern mobile devices does present a unique opportunity for the user to interact with the data. Users with vision loss can use gestures to navigate through the data set and to explore by touch where data points or function plots lie in the tangible context provided by the



(a) (b) (c)

Fig. 4: Interacting with audified displays: drawing a perceived function plot and listening to it (a), interacting with a histogram (b), interacting with a scatterplot (c)

physical boundaries of a touch screen device. Furthermore, the use of physical gestures may aid users in building a mental model of the data.

After the initial sonification pass, the user can explore both discrete and continuous data (Figure 4). The proposed methodology suggests the use of multi-touch gestures to:

- Interactively explore continuous plots by drawing the perceived shape on the touch screen using a single-touch drag gesture, and comparing its audified display with audified display of the original function.
- Interactively explore histograms and scatterplots by touch using a single-touch drag gesture.
- Redisplay data by replaying the initial sonification.
- Zoom using the pinch gesture and reset the zoom factor with a touch-and-hold gesture.

Histograms are explored by sliding a finger over a screen. The screen is treated as a set of vertical virtual keys (Figure 4b). The number of keys is equal to the number of bars in a histogram. When a user touches a key, a sound is played. Pitch scaling and 3D positioning of this sound correspond to the height of the bar that the key represents (Figure 4b). By dragging a finder from left to right (or vice versa), the user can produce a melody-like sequence of sounds, in effect "playing" the data.

For scatterplots, a sound is played whenever the user's finger touches a data point (Figure 4c). This allows the user to investigate data by touch and to trace the overall outline of the data, once again "playing" the data and creating a more accurate mental model of it. Scatterplots are intrinsically more difficult to sonify than histograms during the initial, non-interactive pass and interactive exploration therefore becomes a necessary tool.

For audified displays of univariate functions, the use of the above methodology would result in exploring thin, invisible lines that take up a very small portion of the overall touch screen. Therefore, a different approach is suggested. The user can draw a shape on the touch screen, which is then sonified for the user. By creating shapes whose sonification most closely matches the original sonified function, the user creates a more accurate mental model of the function (Figure 4a).

Replaying the sonification is triggered by a double-tap gesture, to avoid confusion with all the other gestures. Voice-activated redisplay is also possible, especially with the current accessibility features built into modern mobile operating systems. However, voice commands are not being considered for this approach, due to their potential disruptiveness in the classroom setting.

Zooming is used only for univariate functions and affects scaling along the y-axis. The width of the y interval is scaled using the zoom factor, accentuating variations in data along the y-axis. The position of the listener remains the centre of the audified space and listener's orientation is unchanged. Zooming is restricted so that function minima and maxima are always present in the audified context and drawn on screen. In other words, the user is prevented from over-zooming and cutting off portions of the graph by the touch screen border. Zooming is accompanied by synthesized voice prompts and additional audio messages are used to indicate when zooming limits are reached. Zooming is reset to the default aspect ratio by a touch-and-hold gesture.

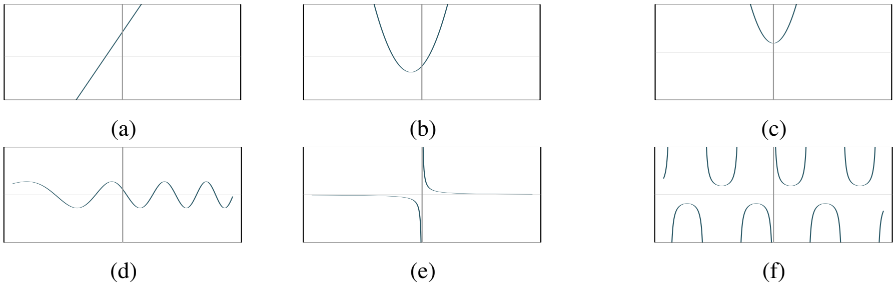


Fig. 5: Continuous test functions: (a) $f_1(x) = 2x + 3$, (b) $f_2(x) = 3/4(x + 1)^2 - 2$, (c) $f_3(x) = x^2 + 1$, (d) $f_4 = 1.5 \sin((0.2x + 3)^2)$, (e) $f_5(x) = 1/x$, (f) $f_6(x) = 1/ \sin x$

3. Evaluation

3.1. Evaluation approach

Previous research has indicated that perception of audio and haptic graphs varies between congenitally blind (blind from birth) and adventitiously blind (late blind) individuals [10]. Prior studies on audio and haptic displays [7] and touch interfaces for people with visual impairments and loss [11] have suggested using blindfolded sighted participants for evaluation. Judgment of blindfolded sighted participants is expected to be similar to that of late blind participants [11]. This is corroborated by a related study of Afonso et al. [13], who investigated spatial cognition of blindfolded and blind individuals in a virtual audio environment. Initial evaluation with sighted, blindfolded users is also useful for calibrating and adjusting the methodology prior to more extensive testing involving participants with vision loss. Therefore the initial testing described below involved blindfolded sighted users.

The proposed approach was implemented and evaluated on an Apple iPad2 using the OpenAL 3D sound engine. To help the participants feel the physical boundaries of the screen, the tablet was equipped with a case that covered the non-display portion of the front panel of the device, and left only the touchscreen uncovered. Six blindfolded sighted participants were used, whose backgrounds ranged from undergraduate degrees in computer science to a PhD in industrial engineering. All evaluators reported being regular, proficient users of touchscreen devices. A qualitative, subjective evaluation was conducted using the think-aloud approach to investigate the users' subjective perceptions of the audio stimuli and the gesture-based interface. To assess how well the mental model created by the sonification procedures matched the original visual representation, for each function and data set, participants were asked to draw the perceived plot. For the scatterplots, users were offered the option of drawing the overall shape of the point spread. For every test case, after drawing the perceived data plot, users were then shown the actual plot. They utilized a Likert-type scale ranging from 1 to 5 to rate how closely their mental model matched the actual plot.

3.2. Test data

Each of the three sonification modes used a separate data set. Univariate function sonification was evaluated with the six functions also used by Grond et al. [6] (Figure 5). This data set contains three relatively simple polynomial functions $f_1(x) = 2x + 3$ (Figure 5a), $f_2(x) = 3/4(x + 1)^2 - 2$ (Figure 5b), and $f_3(x) = x^2 + 1$ (Figure 5c). The two parabolas (5a and 5c) allow users to hear a similar function with and without x-axis intersection. Function $f_4 = 1.5 \sin((0.2x + 3)^2)$ (Figure 5d) is a continuous function with multiple maxima and minima, which intersects the x axis multiple times. Functions $f_5(x) = 1/x$ (Figure 5e) and $f_6(x) = 1/ \sin x$ (Figure 5f) feature asymptotic discontinuities, contain both positive and negative values, but do not intersect the x axis. They represent cases where there is no audible click to inform the user of the function sign change.

To test histogram and scatterplot sonification, a publicly available smoker data set [14] was used, with the following variables: *age* = age of subject in years, *fev* = forced expiratory volume (FEV) in litres, *ht* = height of subject in cm, *sex* = female, male, and *smoke* = non-smoker, smoker. Test histograms used quantitative variables *age*, *fev*, and *ht*, as well as the categorical variable *smoke*. Test scatterplots used variable pairs *smoke*-*fev*, *age*-*fev*, *height*-*age*, and *fev*-*height*. These histograms and scatterplots are shown in Figure 6.

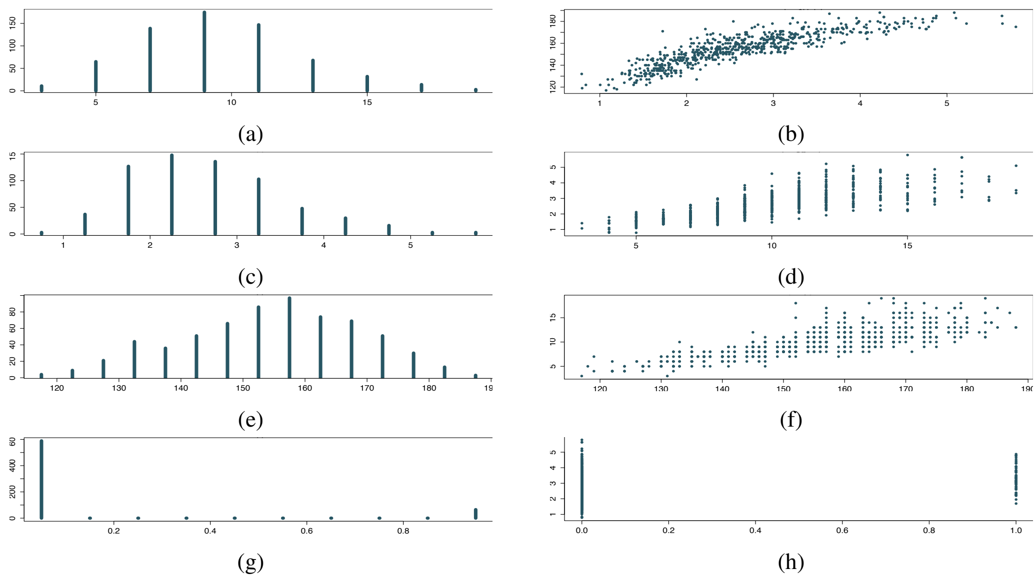


Fig. 6: Histogram test cases, left column: *age* (a), fev (c), *ht* (e), and *smoke* (g); scatterplot test cases (right column): *smoke*-fev (b), *age*-fev (d), *height-age* (f), and fev-*height* (h)

Variables *age*, fev, and *ht* represent easily recognizable histograms for users without only basic university-level mathematical education. All are roughly normal in appearance, but contain minor differences. Histograms of data sets *age* (Figure 6a) and *ht* (Figure 6c) are somewhat skewed to the left, while histogram of fev (Figure 6e) is slightly skewed to the right. Due to the somewhat smaller count of FEV in the bin centered around 135 relative to its two neighbours, the displayed histogram of fev has a somewhat bimodal appearance. It was included in the set to check if users would be able to perceive the shorter bar among several longer bars when the entire sequence is increasing in length (i.e. pitch). The histogram of the categorical *smoke* (Figure 6g) variable only contains two bars and represents an intentionally unintuitive and difficult test case.

Similarly, scatterplots have three similar, simple cases, as well as a less obvious one. The scatterplot of fev-*height* (Figure 6b) has a large number of closely spaced points, while histograms of *age*-fev (Figure 6d) and *height-age* (Figure 6f) have a more ordered visual appearance, particularly *age*-fev due to the small number of possible integer age values. The distinct vertical bars result in a sound representation that audibly different from that of fev-*height*. Finally, *smoke*-fev (Figure 6h) is included as a difficult test case, to determine how the users would perceive and interact an unusual scatterplot - in this case, a somewhat non-sensical correlation of a categorical data set with a quantitative one.

3.3. Results

For each user, the evaluation session took approximately one hour. While this exceeded the expected 30-minute duration, none of the evaluators reported boredom or fatigue. On average, approximately 90% of the time was spent by the evaluators on interactively exploring the data, instead of replaying the initial audification pass.

All evaluators reported being able to create matching models of five out of six continuous functions and rated all matches between 3 (somewhat matching) and 5 (perfectly matching). One test case proved to be an exception to this: $f_2(x) = 2x + 3$. Evaluators did not perceive it as a straight line, and instead reported envisioning a sigmoidal or logarithmic function. The ability to draw an approximation of the audified graph was used by every evaluator. Evaluators reported being able to visualize function plots without the pinch zoom functionality. Some evaluators experienced difficulty executing the double tap gesture blindfolded, despite being proficient touchscreen users.

All evaluators reported being able to create somewhat matching to perfectly matching mental models of bar graphs and found the drag gesture interaction with the bar graph helpful in visualizing the graph.

Evaluators reported having the most difficulty with the scatterplot sonification, particularly the *smoke*-fev scatterplot with two distinct visual clusters. Nevertheless, all evaluators were able to identify two distinct clusters in that test case, although they could not always reliably establish the shape of the clusters or the distance between the clusters. In addition, five out of six evaluators were able to draw models indicating positively correlated data in *age*-fev, *height*-*age*, and fev-*height*. Due to the small sample size, it was not possible to establish if there was a relationship between the evaluators' statistical experience and perception of scatterplots. However, all evaluators found the drag gesture interaction with the scatterplot helpful in visualizing the shape of the point cloud.

4. Discussion and conclusion

These results indicate that the interactivity provided by a touchscreen device enhances utility of sonification displays. Evaluators were able to build approximate, but broadly accurate mental models of the plots that were presented to them. In addition, based on the recorded timings and post-test interviews, evaluators preferred interactive data exploration through multi-touch gestures to passive listening to audified data. Furthermore, evaluators reported becoming progressively more comfortable with the interactive audification software. This suggests that all further evaluation must include a standardized training period for all participants.

Follow-up studies will involve comparisons of the proposed methodology to existing, non-interactive methods using quantitative metrics and larger sets of subjects. They will also investigate new interaction options and gestures by groups of blindfolded sighted users. Additional evaluation will be conducted in collaboration with several educational institutions and involve both congenitally blind and adventitiously blind individuals. In addition, further investigation into the use of HRTFs and binaural audio to aid spatial sound localization in audio displays will be conducted.

Overall, the results show that mobile touchscreen devices are promising platforms for assisting users with vision loss in working with graphs and visualized data. These devices are small, portable, and work wirelessly, thus giving students full-time access to auditory displays. This is expected to greatly improve students' ability to study independently. This may also improve social integration of visually impaired learners and help them participate in group learning activities.

References

1. The cost of vision loss in canada: Summary report. Tech. Rep.; CNIB; 2009.
2. Hermann, T., Hunt, A.. *The sonification handbook*. Berlin: Logos Verlag; 2011.
3. Mansur, D.L., Blattner, M.M., Joy, K.I.. Sound graphs: A numerical data analysis method for the blind. *Journal of Medical Systems* 1985; **9**(3):163–174.
4. Edwards, A.D.N., Stevens, R.D.. Mathematical representations: Graphs, curves and formulas. In: Burger, D., Sperandio, J., editors. *Non-Visual Human-Computer Interactions: Prospects for the visually handicapped (Proceedings of the INSERM Seminar Non-visual presentations of data in human-computer interactions, Paris, March 1993)*. 1993, p. 181–194.
5. Brown, L.M., Brewster, S.A., Ramloll, R., Burton, M., Riedel, B.. Design guidelines for audio presentation of graphs and tables. In: *Proceedings of the 2003 International Conference on Auditory Display*; ICAD2003. 2003, p. 284–287.
6. Grond, F., Drossard, T., Hermann, T.. SonicFunction: experiments with a function browser for the visually impaired. In: *Proceedings of the 16th International Conference on Auditory Display*; ICAD2010. Washington, D.C., USA; 2010, p. 15–21.
7. Riedenklau, E., Hermann, T., Ritter, H.. Tangible active objects and interactive sonification as a scatter plot alternative for the visually impaired. In: *Proceedings of the 16th International Conference on Auditory Display*; ICAD2010. Washington, D.C., USA; 2010, p. 1–7.
8. An introduction to interactive sonification (guest editors introduction). *EEE MultiMedia* 2005;**12**(2):20–24.
9. Hermann, T., Ritter, H.. Listen to your data: Model-based sonification for data analysis. In: *Int. Inst. for Advanced Studies in System research and cybernetics*. 1999, p. 189–194.
10. Yu, W., Ramloll, R., Brewster, S.. Haptic graphs for blind computer users. In: Brewster, S., Murray-Smith, R., editors. *Haptic Human-Computer Interaction*; no. 2058 in Lecture Notes in Computer Science. Springer Berlin Heidelberg; 2001, p. 41–51.
11. McGookin, D., Brewster, S., Jiang, W.. Investigating touchscreen accessibility for people with visual impairments. In: *Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges*; NordiCHI '08. New York, NY, USA: ACM; 2008, p. 298–307.
12. Kane, S.K., Wobbrock, J.O., Ladner, R.E.. Usable gestures for blind people: Understanding preference and performance. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; CHI '11. New York, NY, USA: ACM; 2011, p. 413–422.
13. Amandine Afonso, B.F.K.. A study of spatial cognition in an immersive virtual audio environment: Comparing blind and blindfolded individuals. In: *Proceedings of the 2005 International Conference on Auditory Display*; ICAD2005. Limerick, Ireland; 2005, p. 228–235.
14. Rosner, B.. *Fundamentals of Biostatistics, 5th Ed.* Pacific Grove, CA: Cengage Learning; 1999.