From Tactile to Virtual: Using a Smartwatch to Improve Spatial Map Exploration for Visually Impaired Users

Sandra Bardot[1], Marcos Serrano[1], Christophe Jouffrais[2]

[1] University of Toulouse – IRIT
Toulouse, France
{first_name.last_name}@irit.fr

[2] CNRS - IRIT
Toulouse, France
{first_name.last_name}@irit.fr

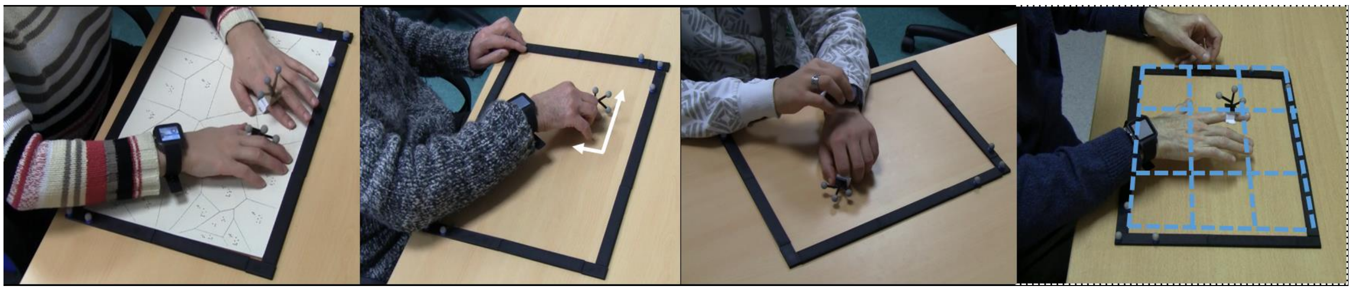

Figure 1. From left to right: exploration of a raised-line map with an infrared tracker on the finger; exploration of a virtual map based on hand tracking and smartwatch feedback; left/right swipes on the smartwatch to filter the data; and raising hands to gather the data within a virtual grid layout.

ABSTRACT

Tactile raised-line maps are paper maps widely used by visually impaired people. We designed a mobile technique, based on hand tracking and a smartwatch, in order to leverage pervasive access to virtual maps. We use the smartwatch to render localized text-to-speech and vibratory feedback during hand exploration, but also to provide filtering functions activated by swipe gestures. We conducted a first study to compare the usability of a raised-line map with three virtual maps (plain, with filter, with filter and grid). The results show that virtual maps are usable, and that adding a filter, or a filter and a grid, significantly speeds up data exploration and selection. The results of a following case study showed that visually impaired users were able to achieve a complex task with the device, i.e. finding spatial correlations between two sets of data.

Author Keywords

Visually impaired users; wearable devices; map exploration; geospatial data.

ACM Classification Keywords

H.5.2. Information interfaces and presentation: User interfaces.

MobileHCI'16, September 06-09, 2016, Florence, Italy
© 2016 ACM. ISBN 978-1-4503-4408-1/16/09…$15.00
DOI: http://dx.doi.org/10.1145/2935334.2935342

INTRODUCTION

Visually impaired (VI) people need regular access to geospatial information during education but also during everyday life. Tactile raised-lines maps are the most commonly used tools for that. However, they are difficult to make, they are expensive, and they depend on the intervention of a tactile graphics specialist. Because raised lines elements and legends are cumbersome, tactile maps cannot contain many details. In addition, they must be edited and printed again when an update is necessary. They generate cognitive issues because Braille legend is usually situated outside of the map, and the reader must interrupt the exploration to read it. Finally, according to the National Federation of the Blind, the Braille literacy rate among blind people in 2009 was reported to be near ten percent. To pass along these limitations, recent prototypes of interactive tactile maps combine a tablet with a tactile overlay [3]. However, the tactile overlay must still be printed and cannot easily be modified, thus limiting the benefits of interactive audio feedback.

Delogu and colleagues explored non-visual access to digital geospatial data relying on the sonification of multiple views, including tables and maps [6]. This work provided an interesting framework for data sonification and manipulation, such as brushing or filtering, but was mainly based on a discrete (keyboard-based) exploration of the data. Especially, they did not study a direct spatial exploration of the map, which may provide the user with accurate hand position awareness on that map.

In this paper, we propose a technique to support direct spatial exploration of data on virtual maps, without any overlay, as a mean to replace raised-line maps (see Figure 1). Our technique relies on the combined use of hand tracking and a smartwatch. The content of the virtual map is rendered using the smartwatch text-to-speech (TTS) engine, for example, a region's name and corresponding unemployment rate, and its vibratory functions to render region borders. The smartwatch is also used as an input device to filter the data by the mean of simple swipe gestures on its touchscreen.

We designed three versions of this exploration technique called Plain, Filter and Grid-Filter. With the Plain technique the watch directly renders the underlying region. With the Filter, the user can select a target data on the watch, which will be the only data to be rendered. With the Grid-Filter technique, the Filter is preserved but, in addition, a spatial layout with nine cells gathers the data. While the Plain and Filter techniques are performed on a surface, the Grid is rendered above the surface and relies on mid-air gestures. When the hand is on the surface, the grid is deactivated to allow precise spatial exploration. When the hand is raised over the surface, the grid is activated to allow a rapid cell-by-cell exploration. Using this technique, we investigated the interest of a rapid but coarse-grained access to the map based on mid-air gestures.

We conducted two studies on these techniques: a comparative experiment and a case study. The comparative experiment involved 12 VI participants, and aimed to compare the usability of the four techniques (regular raised-line maps vs. virtual maps using the three techniques) to access simple geospatial information (name of fictive regions with the name of their main cereal production). Participants explored maps of increasing complexity (increasing number of regions) randomly generated. The second study involved four VI users and aimed at specifying the usage of the Filter and Grid-Filter techniques in scenario including two datasets to be compared (i.e. the population and unemployment rates of the 48 states of the USA). This type of comparison, where the user looks for spatial correlations among multiple types of data, is usual in geo-data visualization (see e.g. [28]).

Our contributions are: 1) a novel solution for VI users to explore virtual maps with data filtering, using a hand tracking technique and a smartwatch, 2) the experimental comparison of the usability of raised-line maps and virtual maps, and 3) a case study relying on a real map exploration with comparison of two sets of data.

RELATED WORK

Our work relates to research on non-visual exploration of maps. We specifically focus on free map exploration as opposed to guided exploration where the user wants to reach a specific point, which involves other interaction techniques (see for instance [13]).

Raised-line interactive maps

As mentioned in the introduction, regular raised-line maps present several major limitations. In order to overcome these limitations, one solution consists in adding a tactile overlay over a touchscreen [3]. Users must double-click identified zones of the tactile map in order to receive audio feedback such as the name and description of points of interest (buildings, parks, streets, etc.) Such a device improves map accessibility because it speeds up learning and it enhances user experience [3]. Furthermore, it provides the user with many layers of information in addition to the name and description of points of interest (e.g. time schedule of public transportation, current movies in a cinema, etc.) However, the tactile overlay must be printed out each time an update is required.

Tangible maps

Using tangible maps instead of raised-line paper maps presents the advantage of adding dynamic haptic feedback, which can be combined with interactivity. McGookin *et al.* [16] designed a device for helping VI people to access graphs with physical objects (phicons) that represent points of the graph. Touching phicons in predefined cells within a tangible grid, the user was able to explore scatter plots or bar charts. However, the system was not designed to explore more complicated spatial representation with numerous points of interest that do not fit in pre-defined cells. They also observed that the objects were regularly knocked over during the exploration.

More recently, Ducasse et al. [7] proposed a novel type of more steady and versatile phicons to construct maps, called Tangible Reels. These phicons are sucker pads that represent points of interest with reels that represent links between pads. Thus, they are specifically suited to represent connections between metro stations or sides in geometric shapes for instance. This device allows VI people to dynamically create maps but also explore existing maps and retrieve specific information related to points and links. Each phicon can be placed anywhere on the screen, and be linked to any other phicon. Hence, this approach overcame the limitation of both raised-line maps (i.e. being static) and tangible maps based on a limited number of positions (see e.g. [16]).

However, tangible maps present major drawbacks. First, the number of phicons that can be used simultaneously is limited because they are cumbersome. In addition, the presence of many phicons (12 being yet large) tends to slow down the hand exploration process [7].

Our main objective was to improve the exploration of geospatial data by visually impaired users. We aimed to design accessible virtual maps that do not rely on any physical artifact (e.g. raided-line map or phicons), but on a mobile device that visually impaired users may already own (e.g. smartphone or smartwatch). Such mobile virtual maps may then be used in many different places (at school or at home for instance).

Virtual maps

Previous work has already explored how VI people can access virtual maps. There are three main approaches for exploring digital maps by visually impaired users: using the keyboard, using haptic devices (e.g. mouse or phantom), and using a touchpad or a tablet.

Using the keyboard, VI people can navigate through maps regions sequentially [18]. The keys are used to move from one region to another one. Another possibility is to divide the map according to a grid layout [28], usually made of 3x3 cells. Users can target a cell with the numeric keypad to quickly get the information related to that cell. These keyboard-based approaches are efficient to navigate a spreadsheet but are less efficient to provide the user with a mental spatial representation of the relative locations of regions on a map. In fact, keyboard-based exploration, being strictly symbolic and discrete, required more cognitive effort for reconstructing the explored layout [6].

Map exploration based on haptic devices generates cognitive issues too. Haptic devices do not provide any stable reference frame, thus differences exist between perceived distances and real distances [12]. In addition, users can lift up the mouse and move it [14], which generates disorientation when operating the mouse without vision [19,24].

Finally, some studies focused on map exploration on a touchscreen or tablet. [25] investigated the sonification of simple navigation maps on handheld touch screens. The evaluation showed that users of the Timbre prototype recognized a few shapes in a matching to sample protocol, but can also develop a rough understanding of an indoor floor plan. [20] asked users to explore a smartphone-based map with vibrotactile and audio output. In that study too, participants correctly perceived basic spatial information. More recently, [9] used a tablet instead of a smartphone, and observed that they were usable but still rely on demanding cognitive processes. In fact, in all these studies relying on touch sensitive screens, it appears that the task of following a line with a finger is difficult, even though the guidance is improved with vibratory and audio feedback. On the contrary, the detection of adjacent zones based on auditory and audio feedback is quite easy. In addition, the memorization of the relative positions of the different zones is effortless because hand exploration is direct (in register with the map). Finally, an obvious drawback of handheld devices is that they only provide a limited surface for exploration, and hence require recurrent panning and zooming operations that are very difficult to perform by visually impaired users.

In the current study, we used a hand tracking technique instead of touch input on a tablet. Hand tracking techniques are now frequent and cheap [15]. They allow multiple hands and fingers tracking, as well as much larger surfaces for exploration than tablets. Furthermore, we used a smartwatch to provide audio and vibrotactile feedback. There are many advantages of the smartwatch over a smartphone or tablet as a personal device. First, with the smartwatch, the two hands are completely free, which is very important for tactile exploration [1]. Then, the smartwatch provides audio and vibrotactile feedback that is collocated in space with the point or the area of interest, which may reinforce the understanding of the map. Finally, the smartwatch provides a second interactive surface that may be used to provide the user with input functions clearly separated from the exploration surface. This segregation of input surfaces may also enhance the general understanding of the map.

NON-VISUAL GEOSPATIAL DATA EXPLORATION

The current work focuses on the spatial exploration of geographic maps with associated data (e.g. demographic or weather maps). Recent work has focused on non-visual techniques to “visualize” complex data (the visualization being tactile or auditory). [28] designed a device relying on multiple views (e.g. spreadsheets or map views) to access demographic data. In that system, users can filter the data within the table, and then navigate the map view with the keyboard. Using that filtering function, they hear the selected data only.

Our work was inspired by this previous study on non-visual data exploration. However, we focused on the complementary objective of improving the spatial component of exploration, providing the user with large surfaces and collocated feedback. We also aimed at providing the user with map access in any situation (home, school, work, etc.) relying on a personal device such as a smartwatch.

Transforming visual maps into virtual maps that make sense for visually impaired users is not straightforward. In the following section, we first analyze the layout and content of regular raised-line maps. We identified the elements that should be preserved to design accessible virtual maps. We then detail the techniques that we designed to explore virtual maps based on hand tracking and a smartwatch.

Tactile raised-line maps for VI people

Tactile raised-line maps have two main advantages: information is tactile, and spatial exploration is direct. They are made according to guidelines (simplification of contours, reduction of the number of elements, legends, etc., see e.g. [26]). These maps contain important elements:

- Contour of areas, rendered through raised lines;
- Points of Interest (POI) and labels represented with specific tactile symbols ;
- a Braille legend describing each area and POI;
- Data associated to each area or POI, for instance the population of a region. In raised-line maps this information is usually written outside the map with Braille.

In addition, tactile exploration of novel spaces relies on behavioral strategies. For instance, [11] has observed specific strategies for the elaboration of object to object relationships. We observed similar behaviors when visually impaired users explore raised-line maps: they frequently come back to previously explored objects or areas; but they also adopt more global strategies, such as scanning the image from left to right and top to bottom. Finally, it is important to note that raised-lines maps support bimanual exploration.

We aimed to preserve these strategies during spatial exploration of virtual maps. We used hand tracking to locate the hands, and we used the audio and vibratory features of the smartwatch to render the information that is under the hand.

Hand tracking for VI people

Spatial exploration by using hand tracking instead of touch input offers several advantages for VI people. First, VI people tend to put multiple fingers and hand palm down on the surface, which, in absence of visual feedback, generates unexpected events [1]. Instead, hand tracking can simply associate one point with each hand or finger. Second, hand tracking allows performing mid-air gestures, for instance to change the information level of the map when raising the hand. Although mid-air exploration may seem difficult for VI people, in this paper, we explore its use to perform coarse-grained spatial exploration (i.e. by using a 3x3 grid). Finally, mobile and low-cost hand tracking solutions have recently been proposed [15], which could leverage making map exploration possible in different contexts and on different surfaces, such as a regular desk, without the need for an underlying touch sensitive surface. Coupling hand tracking with a wearable device for input and feedback (e.g. a smartwatch) makes it possible for VI people to explore maps in many places such as school or at home.

Using a smartwatch for localized feedback

Recent work has shown the interest for VI people to use wearables [27], and particularly smartwatches: they are small, easily accessible and unobtrusive to wear, and can improve information access and social interactions. Wrist is the preferred part of the body for a wearable [21]. Current smartwatches have the advantage of including speakers and vibratory feedback.

We decided to use a smartwatch to leverage hands-free map exploration. We used the smartwatch both as input and output. As input, the device's touchscreen is used to filter or brush data by performing simple horizontal or vertical swipe gestures (Figure 2). As output, the device is used to render localized Text to Speech (TTS), for instance the name of regions. The vibratory feedback is also used to render information, such as the name of POIs or the border between regions. We identified different mappings between these input/output modalities and the map exploration task. We designed three different exploration techniques (Plain, Filter and Grid-Filter) based on the smartwatch.

Plain exploration

Plain exploration is the exploration of a virtual map equivalent to the exploration performed on a raised-line map: each element on the map is rendered.

Input interaction: The smartwatch is only used as an output for this technique.

Feedback: We combined auditory and vibratory feedback. TTS reads out information underneath the hand, such as the name of the region and its population. A 100 ms vibration notifies the transition from one region to another one. A continuous vibration means that the hand is outside the map.

Filter exploration

Filtering data before exploration allows reducing the amount of information to render through TTS, and thus reduces the user's cognitive load. The filtering allows selecting a sub-range of values, for instance regions with more than 100 thousand residents. To perform the filtering, users make swipe gestures on the smartwatch.

Input interaction: A succession of horizontal finger swipes on the smartwatch reads out the filter values (depending on the scenario). A double-tap selects the current filter.

Feedback: After selection, only the data that corresponds to the selected filter is read out. According to the filter state, TTS reads out information underneath the hand, such as the name of the region and its population. As in the Plain mode, a 100 ms vibration notifies the transition from one region to another one. A continuous vibration means that the hand is outside the map.

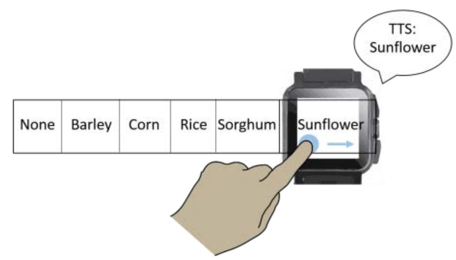

Figure 2. Swipe gesture to select a cereal among others

Grid-Filter exploration

As previously mentioned, filtering reduces exploration time and user's cognitive load. However, it can be difficult to find certain regions in a map especially if they are small. To get a full glance of a map without missing any region, one solution consists in using a 3x3 grid [28], i.e. reading out the information concerning all the regions contained in each cell of the grid.

However, when gathered within a grid, the spatial relationships between regions are masked. To overcome this limitation, we combined the Plain exploration mode with a Grid-based exploration mode. The user can use one or the other interaction level according to hand height above the map.

Input interaction: When the hand is lying onto the table, the user explores the map in Plain mode. When the hand is moving over the table, the user explores the map in Grid mode.

Feedback: At the surface level, the interaction is identical to the aforementioned Filter technique. At the Grid level, a 100 ms vibration notifies the border between two cells of the grid. A double vibration pattern is used to notify the user when he changes the interaction level, i.e. when he is raising or lowering the hand.

COMPARATIVE STUDY: MAP EXPLORATION

The goal of our study was to compare the effectiveness of our virtual map versions against a raised line printed version during an exploration task. We evaluated each technique described above on fictive maps.

Task and instructions

The task was to explore a map, and answer a question as fast as possible. The question was the same for all trials: *Give the name of the four regions that contain <name of a cereal>*. If the participant had not found the regions in less than 4 minutes, we considered the trial as a failure. The participant was allowed to provide each response sequentially (during the exploration) in order to avoid memorization.

Maps

We designed four sets of fictive maps having the same size (A3), but different number of regions: 30, 45 and 60 regions (Figure 3). To create the maps, we used the Voronoi algorithm [8] configured to randomly generate regions of different areas fitting in a 29.7 * 42 cm surface (A3 format). We chose that size because it is the format preferred by visually impaired people [2].

Twenty-seven raised-line maps were printed on A3 sheets of papers. In order to reduce the number of printed maps, we reused raised-line maps with different participants. However, for a given map, we changed the name of the cereal to be found to ensure that we never repeated any condition.

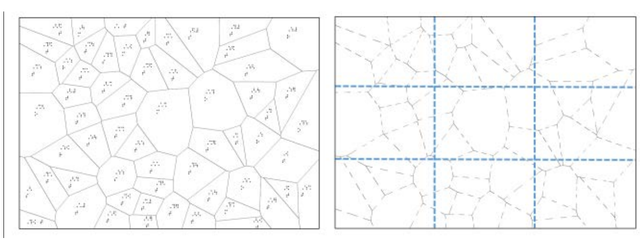

Figure 3. Raised-line (left) and virtual (right) map with 60 regions. Each region includes a number and a letter (braille in the raised-line map). The grid is shown on the virtual map.

Each region within the raised-line maps included two Braille labels corresponding to the region's name and the most common cultivated cereal. To reduce the size of the labels, the name of the region was represented by a number and the cereal by its initial. For instance, the region 11, which contains corn, was described as "11 C".

We generated 324 virtual maps so that we never reused any virtual map across participants and trials: each participant used 27 maps for each virtual technique. Virtual maps present the same general layout and the same labels (number and initial), which were rendered with a TTS engine. We used a 100 ms vibration of the smartwatch to render borders between regions.

Participants

Twelve visually impaired participants (5 females), aged between 20 and 65 years (M=47, SD=13), volunteered for this experiment. All of them were proficient in Braille reading: they declared an average subjective reading proficiency of 2.9 on a scale of 1 to 5 (SD=1.3, min=2, max=5). The level of visual impairments varied: 7 of them were legally blind, and 5 had residual light perceptions (they were blindfolded during the study). Six of them had a bachelor degree, five had a master degree and one was a PhD student. Concerning their occupation, 2 were students, 3 were pensioners, 3 had a work and 4 were unemployed. Eleven of them have a smartphone that they daily use; one has a tablet and none of them had ever used a smartwatch.

Design and procedure

The experiment followed a 4x3 within-participants design with Exploration Technique *(Raised-line, Plain, Filter and Grid-Filter)* and Number of Regions (30, 45, and 60) as factors. The conditions were grouped in blocks including only one Exploration Technique. Within each block, participants repeated three trials for each Number of Regions. We counterbalanced the Number of Regions within each block. The whole order of blocks and maps was counter-balanced across participants. We informed users that they could take a break between blocks. Before using a technique, participants completed a training session. During the training session, we asked users to definitely choose one hand to perform the exploration, and the other one to interact with the smartwatch. They were told how to use the technique, and they were asked to find some regions with associated data. Once they felt comfortable with the technique, they could start the block. All participants chose to use their non-dominant hand to explore and wear the watch, and their dominant hand to perform swipe gestures on the watch.

Apparatus

For hand tracking, we used infrared optical markers tracked by 8 OptiTrack cameras (1mm precision). The system senses the 3D position of markers (x, y and z) at 100HZ. Markers were positioned on the index finger of each hand and on the corners of the interaction surface (Figure 4). We used an Android smartwatch SimValley AW-414 (91 grams, 45.3*44.3*14.1 mm, 28*28 mm touchscreen) with Google TTS. We used TCP sockets over a local Wi-Fi network to connect the watch and the cameras to the main computer.

To set the different parameters of the smartwatch (TTS speed and volume), we carried several user testing. The speed of the TTS engine was twice faster than the default Google TTS speed. Vibratory feedback was set to 100 ms long when crossing a border between two regions or two cells within the grid. It was set to two pulses of 100 ms when raising the hand up or down, and thus changing of exploration level (Plain to Grid level and vice-versa). A continuous vibration indicated that the hand was out of the map.

Figure 4. Left: Experimental setup with the infrared cameras. Right: marker on the finger.

Concerning interaction with the watch, we used the default Android *onFling* callback to detect swipe gestures, and the *onDoubleTap* callback to detect double taps. We defined distance and velocity parameters based on user testing to ensure that swipes could be easily performed.

Collected data

We logged all tracking data (hand movements). For each trial, we measured the completion time as follows: we started the timer when the user had understood the question and was ready to explore the map; we ended the timer when the user had answered the question.

After each block, participants had to fill a NASA-TLX questionnaire [10] about the technique that was just used. At the end of the session, we collected users' preference, and the aspects they liked and disliked about each technique. We also asked whether they used exploration patterns or strategies.

We collected 4 Techniques x 3 Number of Regions x 3 repetitions x 12 participants = 432 trials.

Results

We used a Shapiro-Wilk test to determine the normality of the distributions of collected data. Because the distributions were not normal and could not be normalized, we used non-parametric Wilcoxon and Friedman tests for two or multiple comparisons respectively. We used the Bonferroni correction when needed ($p<.008$ for multiple comparisons between conditions).

Time performance

A Friedman test revealed a significant effect of the Technique factor on the mean time needed to answer questions ($\chi^2(3)=30$, $p<.01$). More precisely, a series of Wilcoxon tests with correction showed a significant difference between the Grid technique and all other techniques: Raised-Line (Z= -5.15, p <.001); Plain (Z= -5.23, p <.001); and Filter (Z= -4.57, p <.001). We also found a difference between Filter and Plain (Z= -5.18, p <.001). Overall, Grid-Filter was faster than the other techniques: on average, answering a question with the Grid-Filter technique took 40 s, with Filter 83 s, with Raised-line 127 s and with Plain 172 s (Figure 5).

When we analyzed the results according to the Number of Regions, we found no difference between Grid-Filter and Filter. However Grid-Filter was always faster than Raised-Line (30 regions: Z= -2.98; 45 regions: Z= -2.82; 60 regions: Z= -3.05; with p<.01). Grid-Filter was also always faster than Plain (30 regions: Z= -3.05; 45 regions: Z= -3.05; 60 regions: Z= -3.05; with p<.01). The Filter technique was faster than Plain for 30 regions (Z= -2.90, p<.01) and for 60 regions (Z= -3.05, p<.01) (Figure 5).

Response times were longer with Raised-Line and Plain. This is due to the fact that users had to thoroughly explore the map in order to find the targeted region and the associated data to answer a question. On contrary, Filter and Grid-Filter renderings quickly provide access to the answer.

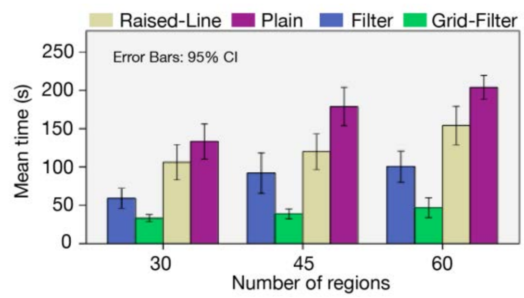

Figure 5. Average time to answer a question for each technique and different numbers of regions.

Precision

Concerning the success rate, i.e. the percentage of regions found, our tests reveal a significant effect of the Technique factor ($\chi^2(3)=25$, $p<.01$). A Wilcoxon test confirmed a difference between Grid-Filter and Raised-Line (Z= -3.29, p= 0.02), Grid-Filter and Plain (Z= -4.62, p<.01), and Filter and Plain (Z= -4.37, p<.01). On average, success rate was 93.1% with Raised-Line, 87.8% with Plain, 98.1% with Filter, and 99.7% with Grid-Filter (Figure 6).

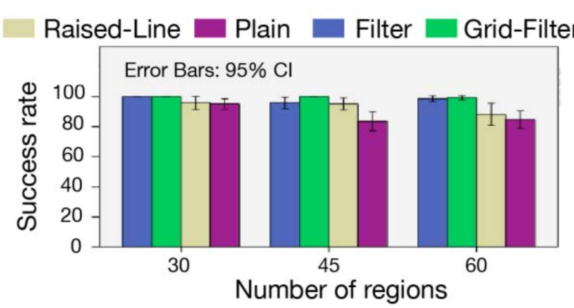

Figure 6. Percentage of correct responses for each technique and different numbers of regions.

For 30 Regions, we only found a difference between Filter and Plain (Z= -2.20, p= 0.02). For 45 Regions, Plain was more efficient than all other techniques: Grid-Filter (Z= -

3.05, p<.01), Raised-line (Z= -2.71, p=0.03), and Filter (Z= -2.66, p=0.01). For 60 Regions, we found a difference between Filter and Plain (Z= -2.80, p<.01) and Plain and Grid-Filter (Z= -2.93, p<.01).

In general, Plain technique was the most difficult technique to perform the task. During the exploration, it was easy to miss several regions. Sometimes users had to browse the map a second time in order to find all regions.

Exploration strategies and hand movements

The observation of hand movements revealed interesting exploration patterns. When exploring the raised-line map, 11 participants used their two hands. With the leading hand, they first read the letter corresponding to the cereal. If needed, they read the number of the region. The other hand was used to find the contours of the neighboring regions to anticipate the following movement (finding the next data).

With the three virtual techniques, most users made either a horizontal or a vertical scanning (Figure 7). Some participants had different behaviors (Figure 7).

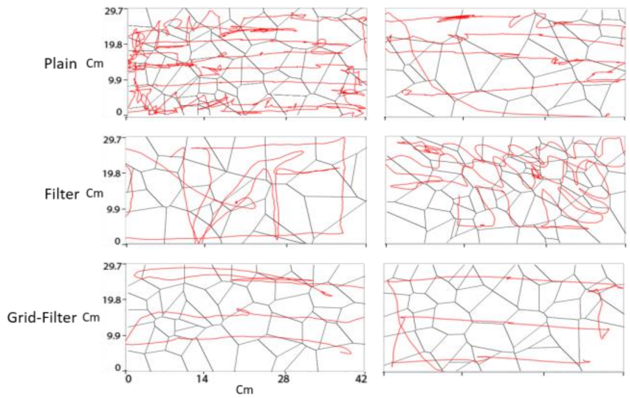

Figure 7. Examples from different participants of hand scanning movement with the three virtual techniques.

NASA-TLX

We found no significant differences among techniques on any of the six NASA-TLX properties. Overall, participants were satisfied with the usability of the techniques: on average, on a 0 to 100 scale (0 being low), participants rated Mental Demand 24, Physical Demand 26, Temporal Demand 36, Effort 26, Performance 70, and Frustration 29.

User preferences

Participants rated the four techniques in order of preference along three criteria: subjective efficiency (which technique is the more efficient?), subjective satisfaction (which one is more pleasant?), and overall preference (if only one, which technique would you use?) Most participants ranked the Filter technique first on the three criteria: 7 out of 12 on efficiency, 6 out of 12 on satisfaction, and 6 out of 12 in general (Figure 8).

Interestingly, the Grid-Filter technique, which was more efficient according to completion time, was only ranked first on efficiency by 1 participant, and satisfaction by 2 participants. It was never ranked first according to the overall preference. One participant reported (P2) that, with this technique, “it was difficult, in mid-air, to estimate his/her own hand location in relation to above map”. P6 reported “it was tiring to keep the hand in mid-air”. However, other participants reported that the “technique is nice because it allows fast exploration” (P5), and because “it gathers information” (P6).

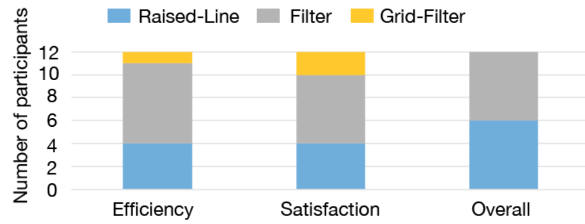

Figure 8. Subjective Efficiency, Satisfaction and Overall ranking.

Concerning the use of a smartwatch, P3 found the swipe technique easy to perform, but, on the contrary, P5 thought it was difficult to use. In fact, P5 is not used to swipe gestures because he has a special phone case that prevents swipe gestures. P7 appreciated “performing hands-free exploration with the watch”. Overall, 11 participants liked using the watch.

Summary

Interestingly, the Grid-Filter technique was the more efficient but not the preferred one. We further explored this technique, compared to the Filter alone in our second experiment.

CASE STUDY: EXPLORING MULTIPLE DATASETS

The goal of this second study was to validate the use of the best techniques (Filter and Grid-Filter) in a more realistic and complex scenario: exploring the map of the USA with two types of data (unemployment rate and density population). Finding data correlation trends is a usual task in spatial data visualization [28].

Map and data

We used a map of the USA with 48 states (we removed Hawaï and Alaska, see Figure 9). This map was not familiar to our participants (average of 2.3 on a scale of 1 to 5). For each state, the user could explore two types of data, unemployment rate and density population. We used two different datasets, one for each technique, from two different years: 1980 and 2010. We used the USA unemployment rate and density population reported by the US Bureau of Labor Statistics and by the United States Census, respectively. For the training, we used a different dataset generated randomly.

Techniques and task

We asked five different questions for each technique, inspired by a previous study on data exploration [28].

1. *(Find Max/Min)* Give all states with a low unemployment rate

2. *(Find a specific state)* Give the population of a certain state
3. *(Compare data)* Among the states with a high population, which ones have a low unemployment rate?
4. *(Find neighbors)* Among the neighbors of the state X, which ones have a high unemployment rate?
5. *(Value in geographical context)* Does population density grow from East to West?

Theoretically, using the Grid-Filter should provide a benefit for the first three questions, as it allows a rapid and synthetic exploration. The question 4 relies on a more precise spatial exploration. Then, the Grid-Filter can help to locate the state but will prevent a correct identification of the neighbors. For the question 5, the Grid-Filter could be used to compare the states in western cells (1, 4, 7) with those in eastern cells (3, 6, 9).

Figure 9. Map of the USA with a 3x3 grid.

Data filtering and user feedback

Data was divided into three types of values: low, medium or high. The TTS engine read out the type of value, for instance: "low". We used different terms for unemployment rate (low, medium, high) and population (small, average, large). For each dataset, the user could select one range of values or all values.

Data selection: Horizontal swipes on the smartwatch selected the data (unemployment rate or population). There were four possible horizontal selections: population, unemployment rate, only state names or all data (Figure 10).

Filtering: vertical swipes selected the range of values for the current data.

Feedback: the TTS engine rendered the name of the state underneath the hand, and then unemployment rate and population density, respectively. Vibrations were as described in the first study (100 ms for changing states or areas and 100 ms twice for changing between exploration levels).

Participants

Four visually impaired women, 47 years old on average, took part in this study. None of them participated in the previous study. Participant 1 (P1), 58 years old, is a Braille teacher and is legally blind. Participant 2 (P2), 56 years old, is a teacher for VI persons and had residual light perception (she was blindfolded for this study). Participant 3 (P3), 19 years old, is a university student and is legally blind. Participant 4 (P4), 55 years old, is a Braille teacher and is legally blind. All of them have a smartphone, which they use daily, and none of them owns a tablet or a smartwatch.

Figure 10. Users can perform left/right swipes to select data (left) and up/down swipes to select values (right).

Procedure and collected data

We counterbalanced the two techniques and the two datasets from different years, so that the (data; technique) couple was specific to each user.

During a preliminary familiarization phase, the participants practiced until they felt comfortable with the technique. Participants had a few minutes to browse the map with only state names rendered to get comfortable with it. During that session, they were shown how to use the filter and/or the grid to retrieve data values concerning specific states. This phase lasted 7 minutes on average.

After the familiarization, they had to answer the first question, without any comment or suggestion from the experimenter. When the strategy used during this first trial was not the optimal one, the experimenter described the optimal strategy. He then asked the user to use it to answer that question once more, but about a different state and data. The same procedure was used for the five types of question mentioned earlier. Hence, each user completed 5 to 10 trials with each technique. We logged the hand movements and we measured the completion time for each trial.

Finally, the participants ranked the two techniques according to subjective efficiency (which one is the more efficient?), subjective satisfaction (which one is more pleasant?), and overall preference (if only one, which technique would you use?) We asked them about the strategy that they eventually used to get the responses. They finally had to mention the aspects they had liked and disliked about the techniques. We also asked them what they would change about the two techniques.

Results

Answers to the questions

All the participants were able to answer the five questions. They performed respectively 13, 11, 10 and 11 trials in total, which means that they were able to find a good strategy on the first trial. On average, they performed 5.7 trials with Filter and 5.5 trials with Grid-Filter to answer the five questions. The only questions where participants were asked to repeat the trial were questions 1 and 3.

On average, they needed 208 s (SD= 68) per question with the Filter, and 148 s (SD=43) per question with the Grid.

Included in that time, the swipe gestures took 39 seconds on average (SD= 20).

Sixty-six percent of the answers were correct with the Filter and 76.2% with the Grid-Filter. Errors only concerned Question 4 (find a neighbor) with the Grid-Filter, because users provided more states than correct.

Data correlation

One user (P3) said that it was convenient to get the two values at the same time, on the same map. She spontaneously compared that condition to raised-line bar charts that she used to explore, and she said that it was more suitable. P1 reported that it was interesting to relate two different types of data on the same map. One of the Braille teachers (P1) thinks filtering on the smartwatch is faster than using a braille map with several data.

Overall strategies with Filter and Grid

Logged hand movements reveal interesting differences between the Filter and Grid-Filter conditions. With the Filter condition, participants browsed the entire map and had to slow down in areas where states were small. However, they still miss some states (see Figure 11). With the grid, users did not miss any state but sometimes had to repeat the TTS feedback. In that case they quickly moved the hand out and back in the cell (see cells 4 and 7 on Figure 11).

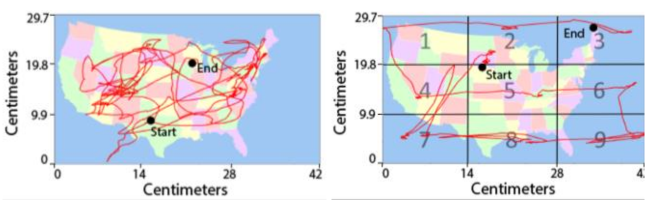

Figure 11. Two examples of logged hand movements for the Question 1 (find max/min) with the Filter (left) and with the Grid-Filter (right).

Two participants (P1, P3) applied the optimal strategies (properly using the two levels of filtering, i.e. data type and value) for all questions except for question 4 (find neighbor) with Grid-Filter: as they used the grid level only, they gave more states than expected. In general, these two participants used systematically the grid level with the Grid-Filter technique. P2 did not like the Filter and used the grid level of the Grid-Filter technique to answer the questions 1, 3 and 4. For the others questions, she browsed the map using only the filter as she already knew some USA states. P4 chose to use only the filter with both techniques: she found it difficult to use the grid level as she felt she could not get a mental representation of the map.

Swipe gestures for selecting and filtering

Most users were able perform horizontal and vertical swipes to filter data. However, we observed that participants had more difficulties performing horizontal than vertical swipes, probably due to the arm orientation that led to diagonal swipes. Three participants (P1, P3, and P4) systematically used the filtering function. P2 never filtered data: she had some difficulties performing swipe gestures, and hence did not like it. She reported that she never uses swipe gestures on her own smartphone.

Using the Grid layout

Although three participants used the grid level, only P3 preferred the Grid-Filter to Filter alone. Users were confident with the Grid-Filter because they found it easy and fast to browse the whole map. However one participant (P4) found that the “grid rendered too much information”. In addition, she did not like it because “it is difficult to get a mental representation of the size of each state”.

Hand-raising and mid-air gestures

Overall, although three of them used mid-air gestures frequently, participants did not appreciate mid-air interaction, which confirms the observations made in the previous study. Three of them suggested replacing the hand raising by another type of gesture or input. Two participants (P2 and P4) said that it was difficult to estimate precisely hand elevation and that they felt uncomfortable. Moreover, most of them reported having difficulties in knowing the precise location of their hand over the map.

Preference

Participants rated the two techniques along three criteria: efficiency, satisfaction and overall. Results were equivalent on efficiency and satisfaction. Two users preferred each technique for each criterion. Overall, most participants (3 out of 4) ranked the Filter technique first.

DISCUSSION AND FUTURE WORK

Raised-lines vs. virtual maps

Our work focused on the usability of virtual maps, including different exploration techniques (Plain, Filter and Grid-Filter), as opposed to the regular raised-line maps. Overall, the results suggest that VI people are able to explore geo-spatial data in virtual maps. More precisely, they show that when filtering functions are added to virtual maps, which is impossible to achieve with tactile maps, they provide the user with an efficient mean to retrieve specific answers about the elements of the map.

Concerning the filtering techniques added to virtual maps, the results showed that Grid-Filter is more efficient than all other techniques, but less preferred. The Filter technique alone appears as a satisfying and efficient compromise. The following case study highlighted that both Filter and Grid-Filter techniques are usable to perform more complex tasks, such as comparing two sets of geo-spatial data over a map. Moreover VI participants appreciated both techniques.

Smartwatches for pervasive access to maps

One of our goals when considering virtual techniques is to leverage map exploration for visually impaired users everywhere (office, school, home, etc.) Our solution is based on hand tracking combined with a smartwatch used to provide the user with both feedback (TTS, vibration) and input to filter data (by type and value). Although it is out of the scope of our paper, embedded and wearable solutions for hand tracking already exist, for instance using a head-mounted camera [5]. Future work will investigate how these solutions adapt to the exploration of geospatial data by VI people.

Two hands for virtual exploration

During our experiment, most participants used two hands to explore raised-line maps. However, our virtual techniques only involved one hand in the exploration task (the other hand was used to interact with the watch before the exploration). Virtual exploration techniques could probably be improved by involving the second hand in the exploration. Most hand tracking systems allow tracking multiple hands. Concerning the feedback, a second smartwatch or just a Bluetooth bracelet with vibration could also be used on the second hand. However, this solution would probably be cumbersome. Another option would be to use vibratory rings, which have the additional advantage of allowing precise feedback on a single finger. In the future, we plan to investigate how to combine multiple vibratory feedbacks on both hands and/or several fingers and, most importantly, how VI people would perceive them. In addition, the vibratory feedback that we provide during exploration could be much richer and rely on a wide range of vibrotactile patterns [17].

Zooming and panning

We have shown that exploration techniques based on virtual maps allow comparing two sets of geospatial data over a map. This task is considered as a complex and useful task. The design space provided by our device is large and allows performing many more tasks that are impossible with raised-line maps, such as zooming and panning. Panning could simply be performed using gestures on the watch or voice commands. Zooming techniques for VI people have already been proposed [23]. They rely on zoom levels with significantly different content, and which preserve the cognitive grouping of information. A simple solution could be to associate a hand gesture with a specific level to zoom in or out.

Mid-air gestures

Our work explored a novel and even provocative approach by proposing VI people to use mid-air gestures. Mid-air gestures are not frequently used in interfaces for VI users. In fact, during design sessions with VI users, it appeared that they are not, at first glance, in favor of these gestures. Touch, including tactile perception on the skin, is the main sensory modality to perceive objects for VI people. Raising the hand away from the object is then not natural. However, our two studies revealed that they are effective to dynamically change the brushing of the geo-spatial data [6], but also the exploration mode that is provided. We observed that all but one participants actively used mid-air exploration when using the Grid-Filter technique, even though they were not required to. In addition, the Grid-Filter technique was the most efficient technique for retrieving specific information from the map. However, many participants reported that it is tiring if it is too long, and that it is difficult to build a mental representation of the map when their hand is moving above the map.

Our studies point out a perceptual issue described by [22]. Indeed, they showed that tactile cues contribute to accurate hand positioning. In the Filter condition of our studies, the VI users explore a virtual map by moving the hand on a surface. They can use both tactile (fingers sliding along the surface) and kinesthetic (arm position and movement) feedback to estimate their own hand position. When they move in mid-air, users must rely on their kinesthetic feedback only (the tactile feedback is missing), resulting in a less precise encoding of hand location in space. However, since mid-air gestures appeared as efficient and useful in our studies, we believe that their use should be further explored. One solution could be to use mid-air feedback, such as Ultrahaptics [4], a system that creates mid-air multitouch haptic feedback.

CONCLUSION

In this paper we proposed virtual spatial map exploration techniques as an alternative to regular raised-line maps. Our techniques are based on the combined use of hand tracking and a smartwatch for feedback and input. We defined three types of map exploration: Plain, Filter and Grid-Filter. In a first study, twelve visually impaired users explored a set of randomly generated maps by using these three techniques as well as the classical raised-line approach. Results show that using the Grid-Filter approach is the fastest, but generates discomfort. In a second study, we observed four VI people who explored two types of data (unemployment and population) on the USA map. Results show that virtual techniques are usable to perform complex tasks such as finding correlations between the two sets of data.

ACKNOWLEDGMENTS

We thank all the visually impaired users who participate to the studies. We also thank the LACII lab as well as IJA and Lestrade institutes, Toulouse, FR. This work was part of the AccessiMap project (research grant AccessiMap ANR-14-CE17-0018).

REFERENCES

1. Brock, A. M., Lebaz, S., Oriola, B., Picard, D., Jouffrais, C., and Truillet, P. 2012. Kin' touch: Understanding How Visually Impaired People Explore Tactile Maps. *Conference on Human Factors in Computing Systems - CHI*, ACM, 2471–2476.

2. Brock, A. M., Truillet, P., Oriola, B., and Jouffrais, C. 2010. Usage of multimodal maps for blind people: why and how. *ITS'10: International Conference on Interactive Tabletops and Surfaces*, ACM Press, 247 – 248.

3. Brock, A. M., Truillet, P., Oriola, B., Picard, D., and Jouffrais, C. 2015. Interactivity Improves Usability of Geographic Maps for Visually Impaired People. *Human-Computer Interaction* 30:

156–194.

4. Carter, T., Seah; S.A., Long, B., Drinkwater, B., and Subramanian S. 2013. UltraHaptics. *Proceedings of the 26th annual ACM symposium on User interface software and technology - UIST '13*, ACM Press, 505–514.

5. Colaço, A, Kirmani, A., Yang, H.S., Gong, N., Schmandt, C., and Goyal, V.K. 2013. Mime. *Proceedings of the 26th annual ACM symposium on User interface software and technology - UIST '13*, ACM Press, 227–236.

6. Delogu, F., Palmiero, M., Federici, S., Plaisant, C., Zhao, H., and Belardinelli, O. 2010. Non-visual exploration of geographic maps: Does sonification help? *Disability & Rehabilitation: Assistive Technology* 5, 3: 164–174.

7. Ducasse, J., Macé, M., Serrano, M., and Jouffrais, C. 2016. Tangible Reels: Construction and Exploration of Tangible Maps by Visually Impaired Users. *CHI'16 - To appear*, 2186–2197.

8. Fortune, S. 1987. A sweepline algorithm for Voronoi diagrams. *Algorithmica* 2, 1-4: 153–174.

9. Goncu, C., Madugalla, A., Marinai, S., and Marriott, K. 2015. Accessible On-Line Floor Plans. *World Wide Web Conference Comittee (IW3C2)*, 388–398.

10. Hart, S.G., and Staveland, L.E. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (eds.). Elsevier, 139–183.

11. Hill, E.W., and Rieser, J.J. 1993. How persons with visual impairments explore novel spaces : Strategies of good and poor performers. *Journal of Visual Impairment and Blindness* 87: 8–15.

12. Jetter, H., Leifert, S., Gerken, J., Schubert, S., and Reiterer, H. 2012. Does (multi-)touch aid users' spatial memory and navigation in "panning" and in "zooming & panning" UIs? *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI '12*, ACM Press, 83–90.

13. Kane, S.K, Morris, M. R., Perkins, A. Z., Wigdor, D., Ladner, R. E., and Wobbrock, J. O. 2011. Access overlays: Improving non-visual access to large touch screens for blind users. *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*, ACM Press, 273–282.

14. Lawrence, M. M., Martinelli, N., and Nehmer, R. 2009. A Haptic Soundscape Map of the University of Oregon. *Journal of Maps* 5, 1: 19–29.

15. Marin, G., Dominio, F., and Zanuttigh, P. 2014. Hand gesture recognition with leap motion and kinect devices. *2014 IEEE International Conference on Image Processing (ICIP)*, IEEE, 1565–1569.

16. McGookin, D., Robertson, E., and Brewster, S. 2010. Clutching at Straws: Using Tangible Interaction to Provide Non-Visual Access to Graphs. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, ACM Press, 1715–1724.

17. Paneels, S., Anastassova, M., Strachan, S., Van, S. P., Sivacoumarane, S., and Bolzmacher, C. 2013. What's around me? Multi-actuator haptic feedback on the wrist. *2013 World Haptics Conference (WHC)*, IEEE, 407–412.

18. Parente, P., and Bishop, G. 2003. BATS : The Blind Audio Tactile Mapping System. *Proceedings of ACM South Eastern Conference*, ACM Press.

19. Pietrzak, T., Crossan, A., Brewster, S. A., Martin, B., and Pecci, I. 2009. Creating usable pin array tactons for non-visual information. *IEEE Transactions on Haptics* 2, 2: 61–72.

20. Poppinga, B., Magnusson C., Pielot, M., and Rassmus-Gröhn, K. 2011. TouchOver map: Audio-Tactile Exploration of Interactive Maps. *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '11*, ACM Press, 545–550.

21. Profita, H. P., Clawson, J., Gilliland, S., et al. 2013. Don't mind me touching my wrist. *Proceedings of the 17th International symposium on wearable computers - ISWC '13*, ACM Press, 89–96.

22. Rao, A., and Gordon, A. 2001. Contribution of tactile information to accuracy in pointing movements. *Experimental Brain Research* 138, 4: 438–445.

23. Rastogi, R., and Pawluk, D. T.V. 2010. Automatic, intuitive zooming for people who are blind or visually impaired. *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '10*, ACM Press, 239–240.

24. Rice, M. T., Jacobson, R. D., Golledge, R. G., and Jones, D. 2005. Design considerations for haptic and auditory map interfaces. *Cartography and Geographic Information Science* 32, 4: 381–391.

25. Su, J., Rosenzweig, A., Goel, A. , Lara, E. de , and Truong, K. N. 2010. Timbremap: Enabling the Visually-Impaired to Use Maps on Touch-Enabled Devices. *Proceedings of the 12th international*

conference on Human computer interaction with mobile devices and services - MobileHCI '10, ACM Press, 17.

26. The Braille Authority of North America. 2010. Guidelines and Standards for Tactile Graphics. Retrieved from http://brailleauthority.org/tg/web-manual/

27. Ye, H., Malu, M., Oh, U., and Findlater, L. 2014. Current and future mobile and wearable device use by people with visual impairments. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, ACM Press, 3123–3132.

28. Zhao, H., Plaisant, C., Shneiderman, B., and Lazar, J. 2008. Data Sonification for Users with Visual Impairment. *ACM Transactions on Computer-Human Interaction* 15, 1: 1–28.