Multimodal Representation of Complex Spatial Data

Hrishikesh Rao
University of Michigan
Ann Arbor, Michigan
hrishir@umich.edu

Sile O'Modhrain
University of Michigan
Ann Arbor, Michigan
sileo@umich.edu

ABSTRACT

For blind users, spatial information is often presented in non-spatial form such as electronic speech. We explore the possibility of representing spatial data on refreshable tactile graphic displays in combination with audio feedback and utilizing both static and dynamic tactile information. We describe an implementation of a New York Times style crossword puzzle, providing interactions to query location and stored data, ask for clues in across and down directions, edit and fill in the crossword puzzle using a Perkins style braille keyboard or a typewriter style keyboard , and verify answers. Through our demonstration, we explore tradeoffs related to available tactile real estate, and overcrowding of the tactile image with a view toward reducing the cognitive workload involved in retaining a working mental model of the active grid, and the time to complete a letter placement task.

CCS CONCEPTS

• **Human-centered computing** → **Haptic devices**; **Accessibility design and evaluation methods**; **Accessibility technologies**.
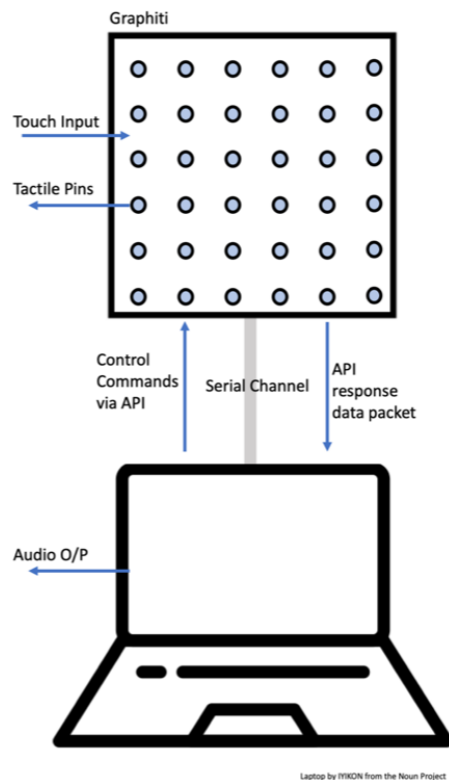
Figure 1: Architecture of the system.

KEYWORDS

Accessibility; Refreshable Braille Displays; HCI; Multimodal Spatial Representation;

INTRODUCTION

Refreshable tactile graphic displays have been in development for some time now and have been used for a number of applications involving spatial representations like maps [5] , GUIs [4], e-readers [3] ,and Sudoku [2]. The large array of controllable pins makes these displays ideal for spatially represented data, however they are limited by the cost of producing very large arrays and by the challenge of packing mechanical actuators tightly enough to achieve high resolution.

Existing approaches to the display of information in non-visual formats typically assume that a blind computer user will choose to use refreshable braille instead of or in addition to the same content being spoken. Here we demonstrate a system that uses a combination of speech and tactile graphics to distribute information for a crossword puzzle, such that the spatial information pertaining to the placement of letters on the grid, and the status of individual grid elements is allocated to the tactile modality while the letters themselves and the word clues are presented using text-to-speech. Our representation of spatial data assists in construction of a mental image of the crossword puzzle through haptic and audio modalities. We also explore the use of dynamic capabilities of the pin matrix as a pseudo third modality in our application. A standard New York Times style crossword puzzle, in its representation, provides various information such as : bounds of the crossword, character locations, clues for words, the direction for the words, and the overall idea of the spatial arrangement of the puzzle. If represented in a standard 6 point braille format like in [2], a 15x15 grid crossword would occupy a space of 45x60 grid dots. The problem with such a large display is its tendency to be confusing, given the size of this representation. With size, the time required to explore the data in order to build a mental image of the puzzle also increases, thus slowing down the user significantly. Our demo targets solving these broad issues and providing guidelines for such a representation.

We are using a refreshable multi-touch tactile display called the Graphiti [1], connected to a computer running MacOS. Graphiti's controllable pin matrix display is twice the size of standard braille dots. The computer is running a python script that communicates with the device over the serial channel using the Graphiti API. Through our work, we aim to open up refreshable braille displays for multimodal applications involving non textual data.

CROSSWORD DESIGN

We used a 15x15 sized New York Times puzzle . The puzzle is bound using a border of pins set at their highest height as illustrated in Fig 2. Every row and column in the crossword were separated by one line of "blank" dots. We tie in a total of 3 closed pins surrounding a given empty box to that box, giving
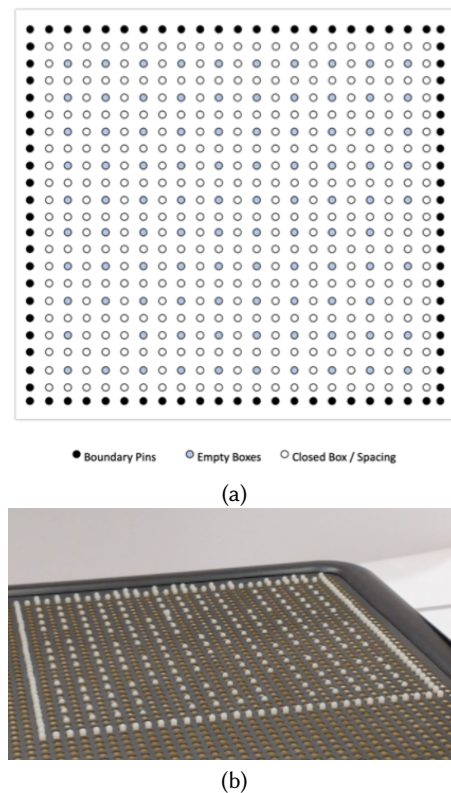
(a)

(b)

Figure 2: Crossword Representation on the Graphiti (a) illustrated, (b) captured during the demonstration.

a total area of 4 dots to trigger a single response. This assists in accurate sensing for all levels of dot height as the distance of the finger from the touch sensitive surface (which changes with dot height) affects the accuracy of detection. Fundamentally, each cell could be in one of three states: Empty, Filled or Blocked. The buffer pins between rows and columns are represented at the same height as Blocked cells. Pins at a higher height tend to be more prominent to users than those at a lower height therefore, we chose a combination that posed the least amount of tactile "noise" among all.

Touch Input: The interaction with the device happens via the touch sensitive surface.

Alert Motion: We have designed a rising and falling pin like motion that occurs approximately twice every second(400ms blinking rate) as a way of bringing a user's attention to a specific location. This event, which we have called the "alert motion", is analogous to a visual blinking cursor, and is integrated with the pins on the Graphiti. This single touch sensitive interface is able to provide the exact row and column often used alongside information related to a specific location (as a pointer) or to mark the position of the current active cell. Coincidentally, the alert motion has a distinct sound produced by the actuators that move the pin at the given location and this is also an additional signal.

Voice Announcements: Interactions like querying for location and clues are exploratory and assist in creation of the mental image of the puzzle. The response to exploratory interactions is auditory announcement of the information such as coordinates, stored character, clues associated with interaction point. Context awareness is an important feature to implement in complex tasks, in order to eliminate redundant steps and improve speed of interaction. Interaction on such a large surface require landmarks to help users orient themselves to the spatial structure of the information presented. As the user fills in the puzzle, the changing status of individual cells serves as landmarks.

Typing: The device is designed for users that are used to screen readers without any requirement for fluency in braille. We have provided the option to use either a Perkins style chorded braille keyboard (Figure 3) or a typewriter style keyboard as our modes of input. As the user types a character on either keyboards, the character typed is announced and the cursor shifts one position. This continues until the user has entered the number of characters equal to the length of the solution, or they exit the edit mode.

DISCUSSION AND RECOMMENDATIONS

This work demonstrates an exploratory step towards multimodal displays that distribute content individual capabilities. In this case, spatial content associated with communicating the layout and status of a crossword puzzle as it is being completed between modalities in accordance with their is allocated to the tactile modality while textual content (clues, words and letters) is spoken. Further, by separating the spatial component from its textual content, the tactile rendering is much more compact, lending itself to low-resolution tactile arrays.

Dynamic Tactile Content: The ability to control the height of the pins on the Graffiti adds another dimension to the display which can be used to draw a user's attention to certain elements. The higher the dot level, the more they 'stand out' for the user passing their hand across the display surface. Hence pins at their highest height are a good barrier to keep users within the bounds of crossword.

Verbosity of Speech Content: We observed that the verbosity of the announced message affected the speed of interaction for users. There is scope to make the announcements more dynamic depending on the expertise and abilities of the user ,like the 'expert mode' in most screen readers.

In summary, through our modality-sensitive implementation of this crossword puzzle demonstration, we have begun to explore how information can be meaningfully distributed to maximize the strengths of the perceptual modalities the user brings to the task. Our next phase of work involves running user studies. We are interested in how such decisions aid in orientation in a complex tactile image, in how they effect a user's interpretation (mental model) of a spatial task, their overall ability to assess own performance and their ability to edit information that is organized spatially. If successful, such design recommendations would be applicable across a range of games that are organized spatially such as Tic tac toe, sudoku as well as applications such as spreadsheets, data visualizations and flowcharts.

ACKNOWLEDGMENTS

REFERENCES

[1] APH. 2016. APH and Orbit Research Announce the World's First Affordable Refreshable Tactile Graphics Display. https://www.aph.org/pr/aph-and-orbit-research-announce-the-worlds-first-affordable-refreshable-tactile-graphics-display/

[2] R Gutschmidt, M Schiewe, and F Zinke. 2010. Haptic Emulation of Games : Haptic Sudoku for the Blind. (2010).

[3] D Prescher, O Nadig, and G Weber. 2010. Reading Braille and tactile ink-print on a planar tactile display. *Lecture Notes in Computer Science* 6180 LNCS, PART 2 (2010), 482–489. https://doi.org/10.1007/978-3-642-14100-3_72

[4] D Prescher, G Weber, and M Spindler. 2010. A tactile windowing system for blind users. *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '10* (2010), 91. https://doi.org/10.1145/1878803.1878821

[5] L Zeng and G Weber. 2012. Building augmented you-are-here maps through collaborative annotations for the visually impaired. *CEUR Workshop Proceedings* 888 (2012), 7–12.

Table 1: Combination of pin heights for crossword representation

Level	Heights	Default Mode
1	background height = 0,empty cells= 2, closed cells = 0, filled cell = 1.	Yes
2	background height = 1, empty cells =3, closed cells = 1, filled cell = 2.	No
3	background height = 2, empty cells =0, closed cells = 2, filled cell = 1.	No

Figure 3: Numbering of the buttons on the Graphiti