AVTMT - The Effect of Multimodal Patterns on Human Memory

Muhammad Tahir, Saleh M. Alshomrani, Abdullah S. Alghamdi
Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, KSA
mtyousaf@kau.edu.sa, sshomrani@kau.edu.sa, aalmalaise@kau.edu.sa

Amina Siddiqua, Maryam Ahmed, Muzafar Khan, Tahir Mustafa Madni, Rizwana Irfan
Department of Computer Science, COMSATS Institute of Information Technology, Islamabad, Pakistan
amina.masood@hotmail.com, mary_ahm@hotmail.com, {muzafar_khan, tahir_mustafa, rizwana}@comsats.edu.pk

Abstract—This paper presents a computer game which consists of multiple tiles at three different levels. When users click on each tile, it flips or un-flips. Each tile, upon flipping, reveals a different patterns e.g. audio, image or text. The system matches those patterns. If two flipped tiles have same pattern, they disappear otherwise they un-flip. This research work 1) examines the effects of audio, visual and text patterns on sighted, deaf, and blind users 2) analyzes the working of human’s memory depending upon the appearance of audio, visual and text patterns and 3) calculates the performance of each user based on the number of previous interactions made with the game. The total number of clicks and the time taken to finish the game was recorded. The results of the four experiments indicate that different types of users behave differently when presented with specific type of information (audio, visual and text).

Index Terms— Human Computer Interaction (HCI), short term memory, spatial memory, performance

I. INTRODUCTION

Human memory is a system that stores, retrieves and processes information received through different senses (sight, hearing, smell, taste and touch). There are different types of memory; Sensory Memory (SM), Short Term Memory (STM), Long Term Memory (LTM) and Spatial Memory (SpM) [1][7] (actually separate memory systems in human brain [14]). SM stores limited information momentarily (like registers in a computer system) while STM provides restricted storage capacity (like RAM in a computer system) and can keep the results of small number of calculations for a short duration of time. STM also supports chunking (remembering information in units like telephone numbers in the unit of 2 digits) [11] which further helps in storing and recalling that information quickly. LTM, on the other hand, has a huge capacity and can store information for a long duration of time. SpM deals with the remembrance of the location of the items in an environment (real or virtual), for example, remembering files and folders positions on different areas of computer desktop. This is the general behavior of human memory for sighted people.

The communication and the working of human memory such as the capability of memorizing objects is somewhat different in blind and deaf people as compared to sighted people [2][3][9]. The hearing sense of blind people is strong as compared to sighted people because former has to focus on the hearing sense in order to interact with others. Same is the case with the sense of vision for deaf users. These three types of users (sighted, blind and deaf) behave different when they are presented with same tasks (visual, auditory, touch) [2][3]. This is probably because of many reasons such as, restricted information processing capabilities (restricted senses), use of different types of input/output devices to perform those tasks and environmental conditions, etc.

The main objectives of our work (AVTMT: Audio-Visual-Text Memory Tiles) include; 1) testing the behavior of different types of users when presented with same tasks 2) observing the performance of SM, STM and SpM of those users for the said tasks (see Figure 3) 3) to test the improvement in performance and 4) to observe the effects of using previous information while playing a game.

II. RELATED WORK

Memory Tiles [10] is a game which consists of multiple tiles, each tile having a unique image. When a player clicks on a tile, it flips and reveals an image. Similarly the player flips second tile, while the first tile is still in un-flipped state. A pair of matching tiles is removed from the playing field and un-matching tiles are flipped back. The author has not tested the game in psychological perspective i.e. it does not provide clues to observe the performance of human memory. Moreover, it is designed specifically for sighted users. In comparison, our work deals with three different types of users i.e. sighted, blind and deaf. It also assesses the behavior of STM, LTM and SpM for storing, memorizing and retrieving information.

A work that targets the users having visual impairments is AudioMemory [13]. It helps to improve
users’ memory by using audio cues. AudioMath, a part of AudioMemory helps in learning mathematics using audio. A comparative study carried out between blind children and children with residual vision (low vision) results that AudioMemory and AudioMath provide an effective way of teaching the related concepts. In another study [16] the authors focused on describing the learning effects of teaching mathematics while keeping in view the different factors including memorizing ability of students. The study [16] focuses on the sighted users while the study [13] highlights the participation of only blind children. Whereas AVTMT emphasizes on blind, sighted and deaf people with different age groups.

Attree et al. [4] discuss how a person remembers an environment and the recognition and location of objects placed in that environment. They investigated the effects of remembering the spatial layout of a virtual environment. They also tested the recognition of objects within that environment by considering the active participation (focused on the interaction with the environment) and passive participation (focused on the tasks) of users. The results indicate that the spatial environment helped active participants to aid spatial memory while passive participants were good at recognizing objects within that environment. Whereas AVTMT also deals with spatial memory but through a non virtual reality game and it does not require specific hardware.

Paul Arnold et al. [12] conducted an experiment to examine the differences in memory of deaf, hearing signers, and hearing non-signers. In this experiment different sets of cards were placed on the table and the user has to find the cards with similar patterns of faces, shoes, and some other objects. The results indicate that the deaf and hearing signers required fewer attempts and recalled the similar patterns quickly as compared to hearing non-signers. Their work is focused specifically on the visual attention while our work provides three different sets of patterns i.e. audio, visual, and text having different faces, audios and objects.

The Gestalt principle of similarity [6][8] states that the objects which are similar to each other in color, size or shape tend to be seen as a unit. By observing this principle in AVTMT, when a user observes similar patterns (visual or text) at different locations on the screen, he or she recalls and groups them logically as one entity.

III. AUDIO-VISUAL-TEXT MEMORY TILES (AVTMT): METHODS AND MATERIALS

AVTMT is a system developed to test the working of human memory using different multimodal patterns. This section describes different aspects of the system.

A. Working

AVTMT is a computer game which deals with the working of human memory especially the STM and SpM. It consists of a board (Graphical User Interface: GUI) which contains different tiles. Figure 1 shows three different tile sets available to users. In the set of visual tiles, each tile contains an image while text tiles contain text and audio tiles contain different sounds. Each game set has three further levels (see Figure 2) of difficulty: easy, medium and difficult containing 16, 36 and 64 tiles respectively.

Figure 1. AVTMT overview (1- visual tiles, 2- text tiles, 3- audio tiles).

AVTMT supports three types of users: sighted, deaf and blind. Sighted users can play all three sets containing three levels each while the deaf users can only play visual and text tiles, each with three levels. Blind users can play audio tiles having easy level only (in our experiments the medium and difficult levels proved to take a lot of time and thus leaving the game in unfinished state by the participants).

When user starts playing the game, the tiles according to the game sets, are displayed. A pattern (image, audio or text) is randomly assigned to each tile. User flips the tile by clicking on it and the corresponding image or text is displayed (played in case of audio pattern). User searches the second tile containing same pattern. System hides (removes) the matching pair of tiles from the board while un-flips the tiles back to original position in case of un-matching tiles.

Figure 2. Different levels of AVTMT (based on the type of game).

When all tiles are removed from the board, the game ends and the results are stored. The total number of clicks a user has made and the total time in which he or she has
played the game is recorded. The performance is also calculated on the basis of number of clicks and the total time taken by a user. This performance further indicates the learning behavior and gradual memory improvement (i.e. by observing the tabular and graphical results).

B. AVTMT Game Types

- **Audio Memory Tiles**: Sighted users can play the audio memory tiles with the help of mouse as input and screen as visual output or speakers as auditory output. Blind users can play audio memory tiles with the help of either microphone or numeric keypad as input and speakers as output.

- **Visual and Text Memory Tiles**: Sighted and deaf users can play these tiles with the help of mouse as input and computer screen as output.

- **Speech Memory Tiles**: These tiles are the variation of audio memory tiles and contain one easy level (16 tiles) (Figure 4). Different sounds are randomly assigned to tiles. Only blind users can play the speech tiles either by using microphone or numeric keypad. By using microphone, users speak the number of the tile to flip it and the corresponding sound is played back. The user then searches the same sound (audio pattern) by speaking the number of another tile. By using numeric keypad, user presses a key from the numeric keypad, the corresponding tile is flipped and a sound is played back. Similarly user searches for the second tile by pressing another key from the keyboard. If the two tiles are matched the system removes them from the board.

B. Pre-defined Variables

Based on the results of very first trials, it was decided to have time factor as a key on the basis of difficulty of the game and the type of user. Moreover the time (named Threshold Time: TT) was reserved in order to provide legacy with existing game [11]. For sighted and deaf users: for easy level, maximum time duration is 1 minute, for medium level it is 2 minutes and 30 seconds and for difficult level it is 6 minutes. For blind users: the time duration is 12 minutes due to their slow task execution (Table 1).

<table>
<thead>
<tr>
<th>Type of User</th>
<th>Audio</th>
<th>Visual</th>
<th>Text</th>
<th>Visual</th>
<th>Text</th>
<th>Audio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sighted Easy</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>720</td>
</tr>
<tr>
<td>Medium</td>
<td>240</td>
<td>150</td>
<td>150</td>
<td>150</td>
<td>150</td>
<td>-</td>
</tr>
<tr>
<td>Difficult</td>
<td>420</td>
<td>360</td>
<td>360</td>
<td>360</td>
<td>360</td>
<td>-</td>
</tr>
</tbody>
</table>

C. Experiments

First experiment was performed for audio tiles. Seven sighted users participated in the audio tiles experiment, out of which 4 of them were female and 3 were male.

Second and third experiments describe visual and text tiles respectively. Total 14 users participated in visual tiles experiment, out of which seven (4 female and 3 male) were sighted and seven (3 female and 4 male) were deaf players. Same users participated in the text tiles experiment. The participants were of 15-25 years of age.

Fourth experiment was performed for speech tiles. Seven blind users, from Al-Maktoom Blind Institute, Islamabad, participated, out of which 3 of them were female and 4 were male. The participants were of 9th and 10th grade students (14-16 years of age).

All users participated in the experiments were familiar to the computing environment and were able to use the

©2013 ACADEMY PUBLISHER
keyboard and other input devices: microphone and numeric keypad.

- **Procedure (for Experiments 1, 2 and 3)**
  The procedure is the same for audio, visual and text tiles. When the game starts, an interface appears that provides options to select user type (sighted, blind or deaf), game type (audio, visual or text) and game level (easy, medium or difficult). After selecting the appropriate options user presses OK button to starts the game. User flips and matches the tiles as explained in section 3.

- **Procedure (for Experiment 4)**
  In speech tiles, an additional option for the input device is provided i.e. to use microphone or numeric keypad. In case of microphone user speaks the tile number and hears the corresponding sound assigned to that tile. While in case of numeric keypad user has to press the tile number and then enter key to play the corresponding sound assigned to that tile. User then searches for the second tile. When two tiles are flipped, the system matches the tiles. If both contain similar pattern then they disappear otherwise they get un-flipped. The game ends when all the tiles disappear.

  The system also provides additional audio feedback in response to user clicks and system state. This helps blind users to easily perform the required actions. This audio feedback is provided when user either correctly matches two tiles or all tiles. For example, user first flips tile ‘1’ and then tile ‘6’ and listen to the corresponding sound patterns. If both tiles contain same sound, they disappear from the board. If the user flips tile ‘10’ instead of ‘6’, both sound patterns are different so both tiles ‘1’ and ‘10’ un-flip. Furthermore, if a tile has already been matched then the system again provides audio feedback such as: tile ‘x’ has already been matched, where x is the specific tile number user is interacting with. The total time user takes to play the game and the number of clicks is all recorded in a database and his/her performance is calculated.

### IV. MEMORY PERFORMANCE CRITERIA

Latest results [5] suggest that the games that are categorized for STM are robustly related to the performance of the memory. While keeping this in view, this study uses two different criteria (describe below) to evaluate the memory performance of a player.

#### A. Total Number of Clicks

As the user finishes playing the game, the total number of clicks made is recorded. Lower number of clicks indicates that user has remembered the position of each pattern assigned to the tile and has easily recalled it. It means users have remembered the images as STM/LTM function and also the location of the image as a function of SpM (see Figure 3).

#### B. Total Time

Every level of the game is assigned a limited threshold time depending on the type of the user and the complexity of the game (as described in the previous section). This time criterion is in accordance with the latest study [15] which results that playing video games improves the encoding speed of visual information in STM and depends upon the time devoted for the game currently being played.

The performance of each user is calculated based on that threshold time. If a user takes time, less than threshold time, it shows that he or she has been able to recall patterns easily (i.e. his/her performance is improved). The following formula has been used to calculate the performance (P):

\[
P = \left(\frac{\text{time taken to play the game}}{\text{threshold time}}\right) \times 100
\]

### V. RESULTS AND DISCUSSIONS

The results are shown in tabular and graphical forms and are based on the time, number of clicks and the performance calculated.

#### A. Results for the Experiments 1, 2 and 3 (Audio, Visual and Text Tiles)

The results of the above mentioned four experiments can further be divided into total of sixteen possibilities based on three game types (audio, visual and text) and three game levels (easy, medium and difficult):

- **Sighted users**: visual easy, visual medium, visual difficult, text easy, text medium and text difficult, audio easy, audio medium and audio difficult (total: 9 possibilities)
- **Deaf users**: visual easy, visual medium, visual difficult, text easy, text medium and text difficult (total: 6 possibilities)
- **Blind users**: audio easy only (total: 1 possibility).

The results for sighted users playing visual tiles indicate that users have taken an average of 40 seconds, 79 seconds and 104 seconds to complete easy, medium and difficult levels respectively. For text tiles their result is 32 seconds, 76 seconds, 101 seconds for easy, medium and difficult respectively. For deaf users playing visual tiles the average time taken in easy, medium and difficult levels was 54, 93 and 121 seconds respectively. While deaf users playing text tiles the result for easy, medium, and difficult tiles was 50, 80, and 113 seconds respectively. All these results are compiled by considering all the participants (including female and males) together. The comparison of average time for female and male participants, separately, on different game types and levels is described in tables II and III.

### TABLE II.

<table>
<thead>
<tr>
<th></th>
<th>Sighted: Audio</th>
<th>Sighted: Visual</th>
<th>Sighted: Text</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Easy</td>
<td>46</td>
<td>53</td>
<td>40</td>
</tr>
<tr>
<td>Medium</td>
<td>203</td>
<td>206</td>
<td>80</td>
</tr>
<tr>
<td>Difficult</td>
<td>377</td>
<td>378</td>
<td>104</td>
</tr>
</tbody>
</table>
TABLE III.

<table>
<thead>
<tr>
<th>Level</th>
<th>Female (Deaf: Visual)</th>
<th>Male (Deaf: Visual)</th>
<th>Female (Deaf: Text)</th>
<th>Male (Deaf: Text)</th>
<th>Female (Blind: Audio)</th>
<th>Male (Blind: Audio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy</td>
<td>54</td>
<td>55</td>
<td>52</td>
<td>48</td>
<td>533</td>
<td>512</td>
</tr>
<tr>
<td>Medium</td>
<td>95</td>
<td>91</td>
<td>78</td>
<td>82</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Difficult</td>
<td>120</td>
<td>123</td>
<td>118</td>
<td>110</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Blind users find the game interactive and easy to use. They were more comfortable (and took less time) in playing the game through numeric keypad rather than microphone. This is probably due to the reason that they frequently use keyboard rather than microphone at their institute.

In Figure 5, the results for the blind users have been re-calculated depending upon the threshold time of 60 seconds as is the case for sighted users. Figure 5 shows the comparative line graph for sighted and blind users who played audio memory tiles (easy level). Results indicate that blind users have taken less time than sighted users to finish the game. This is probably due to the frequent use of their auditory sense which has been greatly improved by time.

Figure 6 represents the comparison with respect to the number of clicks different users performed. This indicates that the blind users performed more number of clicks than sighted users. Similarly sighted users performed more number of clicks than deaf users. For blind users; this is probably due to the involvement of the microphone and numeric keypad as input rather than using mouse click as was the case with sighted and deaf users. Figure 7 explains the average number of clicks performed for visual tiles. The linear trend line indicates the gradual change in difficulty and it is also reflected by the time taken for each level as we move along x-axis from left to right.

Moreover, other results indicate that sighted user playing visual game (easy level) have a combined performance of 67% of threshold time to complete the game. Similarly 53% and 29% of threshold time was calculated for medium and difficult levels. Low value of percentage indicates increase in performance i.e. users have finished playing the game in less time.
VI. CONCLUSION AND FUTURE WORK

The existing memory games generally provide visual or audio patterns and are mostly for sighted users. We focused on sighted, blind and deaf users in our application. The main objective was to analyze the memory behavior of these users i.e. how easily they can recall and recognize either visual/text or audio patterns and their location on the board. Results show that the users find it fun to play AVTMT game.

This work can be extended by obtaining tactile feedback (using either Braille cells or other haptic devices) as output for blind users. Increasing the number of audio tiles for visually impaired users is another important factor that can be added in this research work. Moreover, this game can be developed into multiplayer network game and can be enhanced by applying the procedure of matching more than two tiles.

In order to provide learning to school children, this game can be made more attractive by adding mathematical calculations behind the tiles. It means by creating two groups of tiles; first group will have simple mathematical expressions, for example 2+2, 6*8, 9/3… and the second group of tiles will have the answers to those questions. Users can explore the results by flipping the appropriate tiles. This would again help in improving their short term and spatial memory.

ACKNOWLEDGMENT

We are thankful to the administration of Al-Maktoom Institute for Blinds, Islamabad, visually impaired/blind and deaf students who volunteered for the experiments, speech therapist Ms. Uzma Jabeen and Ms. Maryam Siddiqua and the Undergraduate Project Committee (UPC) of the Department of Computer Science, CIIT, Islamabad for their support and precious comments in improving our research work.

REFERENCES