Applying Audio Cues to Increase Manga Immersion

Sergei Issaev
The University of British Columbia
sergei740@gmail.com

Maryam Zahiri
The University of British Columbia
maryam.zahiri00@gmail.com

Michael Cao
The University of British Columbia
michaelcao@ece.ubc.ca

Lynx Lu
The University of British Columbia
KevinLynx_L@hotmail.com

ABSTRACT

Manga is a type of Japanese comic book which is widely adopted by readers globally. Popular manga series also contain an animation named anime. Unlike manga which is highly static by nature, readers find anime more immersive because it contains additional dynamic components. Auditory information in anime grabs the user's attention and keeps the user immersed within the scene. This key observation raises one question–would adding auditory components in manga improve the manga reading experience? To answer this question, we built a manga reading application which embeds audio into panels. We selected participants and tested on five conditions: no audio, only background music, only sound effects, combined, and combined with inappropriate audio. Data was collected using questionnaires and semi-structured interviews. Using statistical analysis and data coding, our findings show audio can improve the manga reading experience under certain conditions. Background music was found to have a larger impact to the manga reading experience than sound effects. Lastly, implications of inappropriate caused various adverse effects leading to time waste and frustration. Findings of this study call for new technologies to identify appropriate audio, and further research in investigating the impacts of audio on the manga reading experience.

Author Keywords

Manga; Computer Vision;

INTRODUCTION

Manga is the term used to refer to a Japanese graphic novel, specifically characterized by highly stylized art. It is available for view in the form of a physical book or on a digital screen. Many popular manga series also have an animated adaptation, known as an anime. In such cases when manga and anime versions of the same title are available, many users often find that they have a preference of one media form over the other. Some prefer the manga, for reasons including faster story pace (anime tends to be longer with 'filler episodes' unrelated to the main plot), greater amount of details and story elements, more appreciation for the drawings, and other various personal reasons. Others prefer the anime, in which the addition of background music, sound effects, animation of characters, and moving scenes—which enhance the viewing experience of the audience. In general, the static presentation of manga is perceived to be not as immersive as anime.

This suggests that most of the reasons why anime is preferred over manga is related to its auditory contribution. Sound is known to improve the perceived quality of a visual display, as well as significantly impact feelings of immersion in a narrative [22, 6]. Soundtracks that set the tone and mood of a scene may reinforce what is happening within the image. Additionally, sound effects may also emphasize actions within an image or remind the audience of a detail lesser in visible saliency.

While manga contains many dynamic components such as flexible panel sizes, speech bubbles, and iconographic symbols [2], it remains static by providing only visual content to the reader. Therefore, while keeping its unique visual features, we ask if auditory modality can be applied to improve the manga reader experience. To this end, we formally define the research questions we aim to unravel:

RQ1: Can audio features, such as background music and sound effects, improve the manga reading experience?
RQ2: Are all audio features equivalently important to improving a manga reading experience?
RQ3: What are the implications of inappropriate audio features (i.e., wrong sounds or volume levels) with respect to the manga reading experience?

This paper focuses on the possible improvement of a manga reader's experience through audio features. To test this, we created a web application which coupled background music and sound effects per manga panel. Audio features were labeled as either 'appropriate' or 'inappropriate' additions in terms of their accordance to the context of the scene. Each participant

CPEN 541, Dec 10, 2019, Vancouver, BC

read manga chapters with 1) no sound, 2) background music, 3) sound effects, or combined background music and sound effects, which were either 4) appropriate or 5) inappropriate. Questionnaires were filled out after each chapter to rate the readers' experience.

The following sections of this paper discuss the related work, the prototype design of our web app, experimental design, results, and discussion of our study.

RELATED WORK

Enhancement to the Reader Experience

Research in enhancing reader experience has been explored in two dimensions. One dimension includes rapid serial visual presentation (RSVP), the use of sequentially showing an observer stimuli at one location in their visual field. Goldstein et al. explored the use of adding sound effects to RSVP in text format to evaluate perceived immersion of the reading experience [11]. Their findings show that there was a significant perceived immersion increase with the addition of audio to RSVP. Holenderski has also observed the effects of adding background music to novels for e-books [12]. Their findings show an increase in imagery and attention. Manga contains a larger amount of visual content than plain text. In other words, the perceived immersion in manga by adding audio may differ greatly. Our work differs from previous work by focusing on manga, and investigating whether the manga reading experience can be improved using audio.

Another dimension of enhancements to the reader experience has explored the use of haptics to involve physical senses to the reader. Mangen observed that the move from novels to e-books has reduced physical senses, leading to a decrease in focus and shallow understanding of material [15]. Various solutions have been proposed to improve haptic feedback of e-books. Yannier et al. proposed a glove which provides vibrations associated to story events [24]. Winfield et al. used a vibrating plate on an e-book to provide tactile feedback while turning pages [23]. Lastly, Cingel and Piper reviewed the interactions of haptic devices for reading experiences [5]. While haptics have been used to improve the reading experience, our work differs in that we focus on the audio aspect of the reading experience.

Comic Manipulation

Several studies have investigated techniques for comic manipulation. Cao et al. proposed an approach to automatically fit artwork into particular manga layouts [3]. Their findings demonstrated manga panels are irregularly shaped, and new methods are required for accurate panel extraction. Sequentially, Pang et al. developed a robust method for automatic panel extraction from manga [20]. Kovanen and Aizawa proposed, in parallel, a layered segmentation model used to determine the proper reading order of text bubbles in a manga page [14]. Panel extraction methods and reading order has inspired works to improve the immersiveness of manga reading. One example includes Cao et al.'s work on animating still manga through the use of camera movements by inferring per-panel movements based on visual patterns [4]. While much work has been done to devise techniques for manga manipulation, there is limited work in applying auditory feedback to manga. Our work builds on previous literature, such as panel extraction, to test the effects of adding audio for enhancing the manga reading experience.

Applications of Audio in User Interfaces

Audio can be classified into two types: auditory icons and earcons. Auditory icons are sound cues which represent an object using properties of the sound's source [8]. Earcons are defined as abstract audio messages, in the form of synthetic sounds, that provide feedback to the user [1]. Auditory icons and earcons have been applied to many domains; audio cues improve a user's perception to their environment or provide additional functionality to a system.

Previous studies have observed the effects of audio to improve a user's perception of their environment. One of the earliest applications of audio included SonicFinder [9], a desktop application which mapped real-life sounds to different common interface objects. Their findings found that the audio noises increased feelings of direct manipulation and improved productivity when interacting with the computer. Gaber et al. introduced auditory cues within a soft drink factory for monitoring purposes [10]. Their observations show that users were more attentive to diagnosing problems and monitoring the plant. Audio Aura [19] provided office workers with auditory icons, such as waves for receiving emails, for statuses in the office. On mobile, AudioGPS [13] played earcons to aid users in carrying out locational tasks. Spatial audio beeping noises were used to identify direction while volume dynamics were used to determine distance. Similarly, Nomadic radio, a wearable computation platform which continuously provides speech and spatial auditory icon cues, was proposed to improve a user's peripheral awareness [21]. While our work is similar to previous studies by adding auditory information to enhance the user's environmental perception, we differ by focusing on enhancing user experience rather than user productivity.

Auditory cues have also been added to systems with the intention of providing additional functionality. Brewster et al. proposed the use of earcons on mobile phone menus without the use of additional screen space. One of the most important usages of auditory information for additional functionality includes applications for the visually impaired. Mynatt proposed a methodology to transform graphical user interfaces (GUI) into non-visual auditory interfaces [18]. Morley et al. designed an auditory system to identify links and text on the web for the visually impaired [16]. Lastly, Murphy et al. proposed adding auditory cues and haptic feedback to help localize the visually impaired on a web page [17]. To the best of our knowledge, we have not seen a study on incorporating auditory feedback in manga.

PROTOTYPE DESIGN

The prototype consists of a navigable web page that contains isolated panels of manga. Web applications provide fast distribution of software without the need for installation. We chose to design the system to display isolated panels for two reasons. First, individual panels remove the understanding of reading direction. Japanese manga are drawn to be read

from right to left which may off put readers who are used to reading from left to right. Second, displaying isolated panels allows readers to focus on specific content thereby improving the readers understanding of the material. Each panel has associated sound effects and background music to understand the effects of audio on the manga reading experience. In order to move between different sequential manga panels, the user can use the left and right key to traverse through a chapter as can be seen from the figure below.

Figure 1. User Navigation Through the Application

The overall front-end user interface (UI) is left purposely empty to allow the user to be more focused on the specific panel of manga and not be distracted.

EXPERIMENTAL DESIGN

The experiment design is split into the following sections.

1. Combined Mixed: A mix of proper and improper sound effects combined with background music
2. Combined Appropriate: All proper audio cues combined with background music
3. Only Sound Effects: No background music
4. Only Background Effects: No sound effects
5. No Sound Effects

The selection of manga excerpts is necessary to perform experiments with the above section configurations. We choose several unique excerpts from Boku No Hero Academia, one of the highest grossing manga series in Japan, for several reasons. Popular manga series are known to maintain a high level of quality in their work. Furthermore, selecting excerpts from different parts of the same manga series reduces the impact of immersion due to different stories. If we were, for instance, to use different manga series in our sections, it would be possible that a user might feel more immersed based on the art style rather than additional auditory information. Episodic scenes, such as action scenes with a single protagonist and antagonist, were selected as they require minimal context of the story. Another benefit of action scenes include many sound effects and dynamic story pacing. Lastly, Boku No Hero Academia contains an anime adaptation which allows us to collect appropriate audio directly from the anime itself.

In each section, manga panels that use different background music as well as sound effects are provided to the test-users for feedback. To acquire appropriate background music, we downloaded the official sound track (OST) used in the Boku No Hero Academia anime. While the anime OST's are publicly accessible, sound effects are not openly available and often only available when fans extract the audio from the anime. Appropriate sound effects were obtained from the official anime if we were able to find the effects online. However, if we could not find sound effects from the original anime, four researchers synchronized to find the most appropriate sound they could find from FreeSound [7], a free online sound repository. We follow a similar process in order to collect in-appropriate noises for the study; four researchers collaborate and collect sounds deemed inappropriate through FreeSound [7]. In order to acquire data on a users' experience with selected background music and sound effects, a questionnaire is provided to gain feedback from participants. We further triangulate by performing semi-structured interviews with participants in validate our findings.

Participants

We posted on the University of British Columbia's (UBC) Electrical and Computer Student Society Facebook group page to recruit participants for our study. We made sure to provide the students with a consent form for questionnaires. A pool of participants with a well-rounded distribution in experience with manga, as well as gender, is used as a sample set. In total, ten female and nine male test-users participated in the study. The average age of the participants was 23.53 years old (standard deviation 2.32 years). The youngest participant is 18 years old while the oldest is 27 years old. Most participants (36.8%) had read over 100 manga volumes in their lives, with only one participant never having read any manga in their life. The remaining participants (57.9%) have read between 1-100 manga volumes. Approximately half had never watched or read Boku No Hero Academia (52.6%).

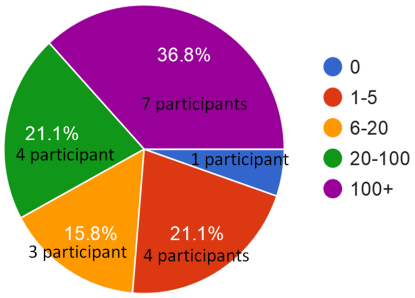

Figure 2. Reading Experience of Participants

Study Procedure

In order to obtain feedback on each experimental condition, the participants were requested to navigate through each of the five different conditions. We requested the participant to complete the condition with no audio first to provide a baseline for the participant's experiment. The remaining four sections were presented in a randomized order for fair comparison. Following each section, participants were asked to complete a

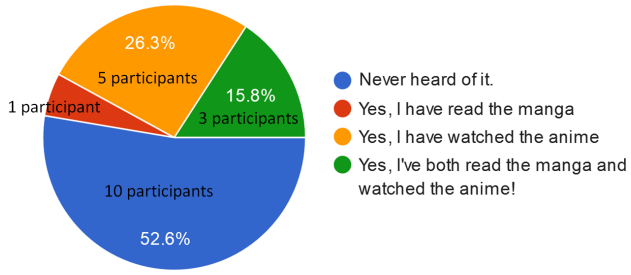

Figure 3. Readers Understanding of Manga

Table 1. Questions asked to each participant after completing each experimental condition

Questions	Scale
What excerpt number are you currently evaluating?	(1 to 5)
I was immersed in the environment of the manga.	(1 to 10) 1 = not immersed at all 5 = neutral (normal manga) 10 = fully immersed)
I felt like I was just perceiving pictures.	(1 to 10) 1 = Strongly agree 5 = neutral 10 = strongly disagree
Compared to reading manga with no additional audio, how would you rate your experience?	(1 to 10) 1 = strongly prefer no-audio 5 = neutral 10 = strongly prefer audio
Overall, I prefer reading manga this way.	(1 to 10) 1 = very negative 5 = neutral 10 = very positive

standard questionnaire with a one-to-ten point scale to rate the amount of immersion, and a one-to-ten point scale whether the participant preferred the current section over no audio.

Interviews

Semi-structured interviews with participants occurred after a participant has performed the study procedure. Usage of interviews enhanced our ability to triangulate and improve validity of the data collected from the questionnaire. The goal of the interview was to gather additional information including: what excerpt was the participant's favorite, which excerpt was most hated, and whether the participant was able to successfully detect the excerpt that had inappropriate sound effects. We further discussed with participants to understand their reasoning behind their selected choices.

RESULTS

Based on the results, it is found that 10 (52.6%) out of 19 participants liked the experiment section that contained both background music and sound effects the most, and 7(36.8%) out of 19 participants liked the experiment section that contains only background music. On the opposite end of the spectrum 2 (10.5%) out of the 19 participants liked having no audio at all. It was apparent that a majority of participants disliked the combined mixed experiment. 14(73.7%) disliked the combined mixed experiment the most, 1 (5.3%) disliked the combined appropriate experiment the most, 2 (10.5%) disliked only sound effects the most, and 2 (10.5%) could not decide. Furthermore, 16 (84.2%) participants were able to identify the combined mixed experiment. The differences in values between a participants preference of a section in comparison to no audio is shown in section 5. Similarly, the differences of values between a participants immersion of a section in comparison to no audio is shown in Figure 5.

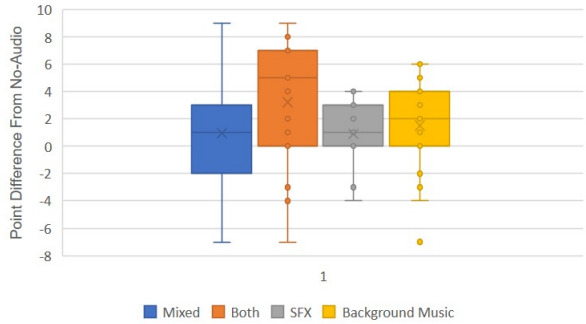

Figure 4. Preference over no audio

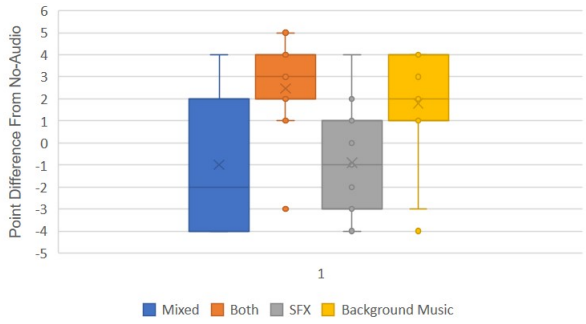

Figure 5. Immersion into Environment

Data coding techniques were used to abstract raw data from interviews into higher level concepts and categories. We answer each research question in more detail, based on the results of the questionnaire and interviews, within the following sections.

RQ1: Does Audio Improve the Reading Experience

As shown in section 5, participants have an overall preference to sections which contained audio over the experimental section with no audio. We believe this observation hints that audio can potentially improve the reading experience, as the user, in most cases, would rather have the audio while reading manga.

Improved reading experience can be inferred from the enhanced immersion of the reading material. With respect to Figure 5, combined appropriate and only background music experimental sections generally provided a positive reader experience. It can also be shown that not all cases in our experiment provided a better reading experience. In particular, the sections with mixed and only sound effect experimental sections were generally shown to decrease the reading experience with respect to the baseline (i.e., no audio). A contrast in immersion with sound effects indicates that the addition

of audio improves the reading experience only under certain conditions.

When coding the interviews, it became apparent of the conditions necessary to improve the manga reading experience. We derived 3 main categories which positively contributed to the manga reading experience: connected audio, atmosphere, and quality. Many participants described that the audio must "connect" with the manga. In other words, it is necessary to select auditory icons which the participant can truly relate back to the objects observed in the manga. Selecting the appropriate audio sounds, in the words of a participant, had "the most emotional investment that went beyond reading a simple manga or cartoon".

It was also described by seven participants that the audio should set the appropriate atmosphere for the manga scene. For instance, one participant described to be confused when the use of bird background music within a gym setting–ultimately reducing their reading experience. Lastly, audio must maintain a certain threshold of sound quality necessary for the participant to have an enhanced manga reading experience. Although one participant admits the sounds correlate to objects in the manga, they described the "[sound quality] wasn't up to standard with expectations". Therefore, it is also important to ensure the quality of selected audio is consistent with the participant's expectations.

Summary: Audio can improve the manga reading experience under certain circumstances. Audio must adequately connect the user to objects represented in the scene and fit in the atmosphere of the scene. Even if the audio fits the scene, sound quality must be on par with the participant's expectations to improve the reading experience.

RQ2: Sound Effects or Background Music

When comparing experiment sections of only background music or sound effects With respect to section 5, it is clear that a user generally would prefer containing either sound effects or background music over no audio. However, when analyzing the same effects on Figure 5, it is quite clear background music has a much higher impact on the reading experience in comparison to sound effects. We suspected this may due to background music providing more emotional context to the manga than sound effects.

While coding data from the interviews, we identified another reason why the immersion of background music was much higher than sound effects. One category includes reading pace. Many readers were found to be restricted by the reading pace due to sound effects. While background music can be spread across multiple panels, sound effects can only be associated with one panel. Fast readers were caught off guard by the sound effects being played at a fast pace, and hence preferred the experiment with only background music.

Summary: Both background music or sound effects are generally preferred over having no audio. Background music is perceived to be more immersive, as it does not dictate the speed of reading the manga in contrast to sound effects.

RQ3: Implications of Inappropriate Audio

While a combination of appropriate sound effects and background music was shown to improve the immersion of the environment (see Figure 5), adding random sound effects displayed significantly lower results. Nine participants rated the immersive experience to be lower than the case with no audio; the remaining ten participants rated, on average, two points higher than the case with no audio. We speculate that half the participants rank the mixed case slightly higher than no audio because, while the mixed case contains in-appropriate audio, the mixed case partially contains appropriate sounds which still immerse the reader.

To answer RQ1, three conditions were described to be necessary to improve the manga reading experience: connected music, atmosphere, and quality. The addition of inappropriate audio caused various negative effects which reduced the manga reading experience. We identified three categories of implications due to inappropriate audio: Confusion, Awkwardness, Frustration. nine participants found inappropriate audio to be confusing. The participants described to be "out of focus" when inappropriate audio was played. Pulling focus wasted many of the participants time by making the participants "confused about what is going on in the [manga]". Pulling focus away from the manga's story line was shown to be the biggest cause of downgraded manga reading experience. Two participants described some audio to be "too cliche", stating the scenes became awkward; the atmosphere of the manga is hindered by the awkwardness audio can produce. one participant was annoyed by the inappropriate sounds, and their disability to not modify the sound provided them with anger. The result of inappropriate sounds led to one participant wanting to "hack the site while [reading the combined mixed experiment]".

Summary: When audio does not represent follow conditions from RQ1, users feel confused, awkward, and/or frustrated. Confusion was the biggest cause for degrading the reading experience.

DISCUSSION AND FUTURE WORK

The original plan for the project involved developing a computer vision model capable of locating, translating, and recognize sound effects in each comic panel. However, due to the highly stylized fonts used to write sound effects, combined with the presence of drawings/shading in the background of the text, no existing computer vision library was capable of effectively capturing sound effects. Tesseract-OCR (Optical character recognition) library was used initially, which failed to pick up any words at all from the manga panels. Next, Google Cloud OCR was used, which picked up most of the text in speech bubbles, but was unable to pick up sound effects.

Current OCR libraries function well when there is a clean, white background behind the text, and the text is written in a font that was seen during the training phase of the model. These assumptions are not met in comic books, since the background of sound effects in panels often contain parts of other drawings, and the font used is manually drawn by an artist. Therefore, it is currently impossible to automate the process of selecting appropriate sound effects for each panel. This will only become possible once an OCR library is

released which has been trained on sound effect-like fonts, as well as on unclean backgrounds. Developing such a library is beyond the scope of this project.

Further work should also be guided towards identification of appropriate and inappropriate audio. As our study finds, sound quality must be on par with the participants expectations to improve the manga reading experience. We leave determining the sound quality, and minimum level of quality required to improve the reader experience for a particular user as future work. Furthermore, unlike anime, which is timed to play audio at exact moments, it is less straight forward when designing a system to play audio under a user's control. Future work can also explore methodologies to handle playing audio such that a user is not startled by additional sounds.

It is interesting to note that no participants in our study selected the only sound effects experimental section as their favorite. In order to continue the investigation, more studies should be done specifically on the sound effects' impact on the reading experience, since no participants voted in favor of reading with only sound effects. However, it is clear that the impact of background music can boost a reader's immersion and allow for a more enjoyable experience. In addition, since the study was performed only on one manga publication, more work should be done in regards to various different genres to obtain a better insight into whether all types of manga can be impacted the same way.

LIMITATIONS AND THREADS TO VALIDITY

Due to the time constraints of this course, we could not evaluate over as many participants as we would have liked. Our sample pool may not be representative of the entire population. Sounds collected for the experiment may not be considered appropriate by the entire population. We mitigate the collection of sounds as best as we can by collaborating together to decide on what sound would be the most suitable. All experimental results are based on a single manga series; it is possible that the manga reading experience may be impacted differently depending on the manga. Participants could have potentially put in random results when filling out their questionnaire. We mitigate this issue as best we can by adding redundant question checks in the form. Furthermore, we triangulate our data with responses from semi-structured interviews.

CONCLUSION

We have studied the overall effects of audio, such as background music and sound effects, on the manga reading experience. It can be concluded that audio has the potential to produce a positive effect in the reading experience of manga, so long as the audio meets certain criteria. Audio must be chosen carefully to ensure the it resonates with users, fits the atmosphere of the story, and maintains quality expected by the user. Our study finds that having only sound effects or background music is preferred over no audio. However, background music has a higher impact to the reader experience because it does not dictate the user's reading speed. Lastly, selecting inappropriate audio can heavily impact the reading experience by causing the reader to be confused, awkward, or frustrated. Side effects of the impact on reader experience mainly cause time waste.

ACKNOWLEDGMENTS

We would like to thank participants who volunteered in conducting the study. Furthermore, we deeply appreciate the guidance provided by Dr. Fels for the project.

REFERENCES

[1] Meera M Blattner, Denise A Sumikawa, and Robert M Greenberg. 1989. Earcons and icons: Their structure and common design principles. *Human–Computer Interaction* 4, 1 (1989), 11–44.

[2] Mio Bryce, Christie Barber, James Kelly, Siris Kunwar, and Amy Plumb. 2010. Manga and anime: Fluidity and hybridity in global imagery. *electronic journal of contemporary japanese studies* (2010).

[3] Ying Cao, Antoni B Chan, and Rynson WH Lau. 2012. Automatic stylistic manga layout. *ACM Transactions on Graphics (TOG)* 31, 6 (2012), 141.

[4] Ying Cao, Xufang Pang, Antoni B Chan, and Rynson WH Lau. 2016. Dynamic manga: Animating still manga via camera movement. *IEEE Transactions on Multimedia* 19, 1 (2016), 160–172.

[5] Drew Cingel and Anne Marie Piper. 2017. How parents engage children in tablet-based reading experiences: An exploration of haptic feedback. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 505–510.

[6] Karen Collins. 2013. Sonic Subjectivity and Auditory Perspective in Ratatouille. *Animation* 8, 3 (2013), 283–299.

[7] Frederic Font, Gerard Roma, and Xavier Serra. 2013. Freesound technical demo. In *Proceedings of the 21st ACM international conference on Multimedia*. ACM, 411–412.

[8] William W Gaver. 1986. Auditory icons: Using sound in computer interfaces. *Human-computer interaction* 2, 2 (1986), 167–177.

[9] William W Gaver. 1989. The SonicFinder: An interface that uses auditory icons. *Human–Computer Interaction* 4, 1 (1989), 67–94.

[10] William W Gaver, Randall B Smith, and Tim O'Shea. 1991. Effective sounds in complex systems: The ARKola simulation. In *CHI*, Vol. 91. 85–90.

[11] Mikael Goldstein, Gustav Öquist, and Staffan Björk. 2002. Evaluating Sonified Rapid Serial Visual Presentation: An immersive reading experience on a mobile device. In *ERCIM Workshop on User Interfaces for All*. Springer, 508–523.

[12] Kacper Holenderski and Jun Hu. 2013. Enriching reading experience with dramatic soundtracks. *Adv. Multimedia Technol* 2, 1 (2013), 24–31.

[13] Simon Holland, David R Morse, and Henrik Gedenryd. 2002. AudioGPS: Spatial audio navigation with a minimal attention interface. *Personal and Ubiquitous computing* 6, 4 (2002), 253–259.

[14] Samu Kovanen and Kiyoharu Aizawa. 2015. A layered method for determining manga text bubble reading order. In *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 4283–4287.

[15] Anne Mangen. 2008. Hypertext fiction reading: haptics and immersion. *Journal of research in reading* 31, 4 (2008), 404–419.

[16] Sarah Morley, Helen Petrie, Anne-Marie O'Neill, and Peter McNally. 1998. Auditory navigation in hyperspace: Design and evaluation of a non-visual hypermedia system for blind users. In *Proceedings of the third international ACM conference on Assistive technologies*. ACM, 100–107.

[17] Emma Murphy, Ravi Kuber, Philip Strain, Graham McAllister, and Wai Yu. 2007. Developing sounds for a multimodal interface: conveying spatial information to visually impaired web users. Georgia Institute of Technology.

[18] Elizabeth D Mynatt. 1997. Transforming graphical interfaces into auditory interfaces for blind users. *Human-Computer Interaction* 12, 1 (1997), 7–45.

[19] Elizabeth D Mynatt, Maribeth Back, Roy Want, Michael Baer, and Jason B Ellis. 1998. Designing audio aura. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 566–573.

[20] Xufang Pang, Ying Cao, Rynson WH Lau, and Antoni B Chan. 2014. A robust panel extraction method for manga. In *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 1125–1128.

[21] Nitin Sawhney and Chris Schmandt. 2000. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM transactions on Computer-Human interaction (TOCHI)* 7, 3 (2000), 353–383.

[22] Russell L Storms. 1998. *Auditory-visual cross-modal perception phenomena*. Technical Report. NAVAL POSTGRADUATE SCHOOL MONTEREY CA.

[23] Laura Winfield, John Glassmire, J Edward Colgate, and Michael Peshkin. 2007. T-pad: Tactile pattern display through variable friction reduction. In *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07)*. IEEE, 421–426.

[24] Nesra Yannier, Ali Israr, Jill Fain Lehman, and Roberta L Klatzky. 2015. FeelSleeve: Haptic feedback to enhance early reading. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1015–1024.