Towards Diverse Media Augmented E-Book Reader Platform

by

Kazi Masudul Alam

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the Master's degree in
Computer Science

Ottawa-Carleton Institute for Computer Science
School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

Abstract

In order to leverage the use of various modalities such as audio-visual-touch in instilling learning behaviour, we present an intuitive approach of annotation based hapto-audio-visual interaction with the traditional digital learning materials such as eBooks. By integrating the traditional home entertainment system and respective media in the user's reading experience combined with haptic interfaces, we examine whether such augmentation of modalities influence the user's reading experience in terms of attention, entertainment and retention. The proposed Haptic E-Book (HE-Book) system leverages the haptic jacket, haptic arm band as well as haptic sofa interfaces to receive haptic emotive signals wirelessly in the form of patterned vibrations of the actuators and expresses the learning material by incorporating audio-video based augmentation in order to pave ways for intimate reading experience in the popular eBook platform. We have designed and developed desktop, mobile/tablet based HE-Book system as well as a semi-automated annotation authoring tool. Our system also supports multimedia based diverse quiz augmentations, which can help in learning tracking. We have conducted quantitative and qualitative tests using the developed prototype systems. We have adopted the indirect objective based performance analysis methodology, which is commonly used for multimedia based learning investigation. The user study shows that, there is a positive tendency of accepting multimodal interactions including haptics with traditional eBook reading experience. Though our limited number of laboratory tests reveal, that haptics can be an influencing media in eBook reading experience, but it requires large scale real life tests to provide a concluding remarks.

Acknowledgements

This thesis is based on research, conducted during my graduate studies at the Multimedia Communications Research Laboratory (MCRLab) and DISCOVER Lab of the University of Ottawa. The work presented here is the accumulated result of highly collaborative efforts and I owe gratitude to the people, I have worked with.

First and foremost, I would like to convey my sincere gratitude to my supervisor, Dr. Abdulmotaleb El Saddik, for giving me the opportunity to pursue research on this rapidly emerging field and move forward along the path of innovation and novelty. His invaluable research ideas and experience kept me motivated all the way towards the progress and completion of this research work. Besides my supervisor, I would like to acknowledge my colleague Abu Saleh Md. Mahfuzur Rahman, for his valuable discussions and effective collaborations. It is worth to mention that my fellow colleagues, Md. Abdur Rahman, Aysha Akther, Faisal Arafsha, Heung-Nam Kim, Md. Anwar Hossain, Md. Shamim Hossain, and SK. Md. Mizanur Rahman were valuable source of information and inspiration throughout my graduate research.

I thank my parents, Kazi Shahidul Alam and Hosneara Jahan, who have been a continuous source of love, support and encouragement all throughout my graduate studies. Last but not the least, I am grateful to my dearest wife, Aysha Akther and our two kids Kazi Tahmeed Alam, and Sameeha Alam, for their love, encouragement and sacrifice to follow my dreams.

Dedication

I dedicate this research to my parents, my loving wife, kids and to the people of Bangladesh and Canada, who supported my studies from the young age up to this higher studies.

Contents

List of Tables

List of Figures

Glossary of Terms

2D Two Dimensional

3D Three Dimensional

API Application Programming Interface

AR Augmented Reality

CG Computer Graphics

E-Book Electronic-BOOK

GUI Graphical User Interface

HDMI High-Definition Multimedia Interface

HMD Head Mounted Display

HTTP Hyper Text Transfer Protocol

IDE Integrated Development Environment

J2SE Java 2 Standard Edition

J2ME Java 2 Micro Edition

LCD Liquid Crystal Display

MCQ Multiple Choice Question

PDF Portable Document Format

SDK Software Development Kit

SL Second Life

SRI Stanford Research Institute

TTS Text-to-Speech

USB Universal Serial Bus

VR Virtual Reality

XML Extensible Markup Language

Chapter 1

Introduction

1.1 Background and Motivation

According to a research article by Harrison [25], French visionaries Robida and Uzanne stated the future of books in 1895. Their prognosis was that written book contents would soon be transferred into sound wave and be played in gramophone and subsequently gramophone would evolve to become pocket-sized player. In contemporary research, MagicBook [8] project was an early attempt to bring augmented reality display in book reading experience, where a user could see 3D virtual models out of the pages. This scheme also supported multiple user immersion in the same scene. In order to replicate physical metaphors used in real books, Chen et al. [15] introduced a new design of dual display for e-book reader and various supported interactions like folding, flipping, and fanning in their system.

Recently, handheld mobile devices are widely adopting tactile feedbacks in their touch screen based interfaces [9]. With the growing popularity of the haptic sensory feedbacks in the mobile handheld devices [26], the challenge is now to find new avenues to intuitively incorporating the feature into popular applications so that interactive storytelling becomes more appealing to the users. To satisfy the demand of ubiquitous computing age, Park et al. [44] suggested a book evolution framework in their survey which identifies three dimensions of book evolution: digitalization, augmentation and hypermediation. They also recommended three research and business development directions: 1) Open Hypermediation, 2) eBook Augmentation, and 3) Mobile-Device Based Augmentation. According to the authors, hypermediation will evolve from internal links to external links and from closed (proprietary) links to open links. They have also shown that, there

will be an evolution from e-book into augmented e-book or augmented book into digitalized augmented book. All of these evolutions are targeted to enhance learning and entertainment experience of the book readers.

Reading is a complex cognitive process, which intends to decode meanings from text by relating words with one existing knowledge-base and understanding. According to Dale [16], more use of modalities ensures effective learning. He described our learning phenomenon numerically 10% of what we read, 20% what we hear, 30% of what we see, 50% of what we hear and see, 70% of what we say and 90% of what we both say and do. Though these numerical labels are not widely accepted but the theory of "Cone of Learning" is mostly acknowledged in the research communities relating to analysis of the reading behaviours.

Our ways of readings are continuously moulded with the advent of new technological innovations, devices and platforms. For example, e-book is continuously challenging the existence of print book and at times e-books are replacing the former in schools [38]. In our daily life, we are getting more and more explored to electronic materials than printed books as now-a-days most of the advanced hand held mobile devices provide e-book reading facilities [57]. Mangent [38] depicted that, *"Haptic perception is of vital importance to reading, and should be duly acknowledged. The reading process and experience of a digital text are greatly affected by the fact that we click and scroll, in contrast to tactilely richer experience when flipping through the pages of a print book"*. As today's user community is widely accepting haptic interfaces; the effect of tactile feedbacks from the reading materials has started to form a new genre of research interest. Currently, many e-book reader devices have introduced tactile haptics as well as sound feedbacks while flipping a page. Haptics is considered to be both as independent and supplementary medium of communication channel [46]. Physical contacts are fundamental needs to mental and psychological development and hence their applications in various cognitive applications have attracted attention of many researchers around the world [24].

Considerable research efforts [8][62][22][11] have been put on to reduce the gap between the real book and virtual reality in the past few years. We find from [44], that several works on augmented books has explored relationship between real and digital books, their various interaction techniques and technologies, audio-visual augmentation methodologies etc. But the limitation of most of the existing researches on augmented books lies in the requirements of specific and unpopular computer interface such as head mounted display, ARToolkit, bar code scanner, projector with cameras etc. As well as, all these research works mostly explore the opportunity of visual immersion let alone

the opportunities of haptic immersion with the e-book materials. This is particularly significant, as tactile (Haptic) feedback has been present in most of the handheld devices now-a-days [26].

Motivated by some of the above works and research opportunities, we present a new e-book reading platform which reflects the direction of Park et al.[44] by combining the concept of hypermedia and e-book augmentation in various platforms such as desktop/handheld devices to evaluate suitability of user's acceptance to tactile feedbacks from e-book reading materials. We have elaborated our system to support broader view of multimedia rather than hypermedia only. We have augmented our e-book reading system with traditional multimedia (audio, video) as well as haptics to analyze the effect of multimodality on immersive reading experience.

1.2 Research Problems

From the contemporary research, we find that most of the research works targeted mixed reality, where virtual reality is augmented with the real book. Many different versions of such systems are being designed and developed where a normal user need to use additional head mounted display or projector with cameras types of devices to enjoy visual immersion. However, there is hardly any work which explains the effect of haptic immersion in reading experience. To the extent of our knowledge, no research article has explored the possibilities of augmenting haptics with eBook reading materials intuitively and explained the effect of such modality in immersive reading experience. As well as, there is no such eBook reading platform which leverages the advantage of a ubiquitous smart space where there is haptic devices (i.e. jacket, sofa, arm band), multimedia entertainment systems (i.e. TV, Sound system). It is worth to be noted that eBook platform is supported in all computing platforms and already considered as next generation of book. Besides, current world is adopting numerous handheld mobile devices (i.e. Smartphone, Tablet) which have the necessary computing capabilities, can also communicate ubiquitously with their surrounding devices. We also need to explore the usefulness of haptics modality combined with already accepted modalities such as text, sound, video etc. for the purpose of immersive and effective reading. Such eBook platform has to be designed such that it can adopt the needs and dimensions of future books. As a result, we need a generic framework for eBook platform which can adopt to multiple modalities, and which can meaningfully convey the information to target devices as well as extended to support additional features used in learning experience tracking.

1.3 Objective and Contributions

Our focus is on ubiquitous environment supported e-book augmentation rather than real book based augmented reality using non-popular visual devices. For example, in a home entertainment scenario (Fig. 1.1), where a father is reading out his travel notes using his iPad, he can choose to show recorded video of the visited places automatically to his wife and the kid, by using the home multimedia screen (e.g. television screen) as well as incorporate them with immersive experience by using haptic enabled devices. Here, haptic sofa[1] and haptic jacket [53] will provide content dependent vibrotactile feedbacks such as boat or train movements. It is also possible to augment such experience with background music playing in the surrounding stereo system. In order to deliver such total (visual, audio, haptic) immersive experience, there is a need to propose Haptic E-Book (HE-Book) which is an e-book reading platform, capable of delivering content related traditional and additional multimedia such as audio-video, vibrotactile feedbacks to user's reading experience. The system architecture should be easily extendable to support other available media such as smell, taste etc. A semi-automatic hapto-audio-visual annotation method is needed in order to support the annotation author. In addition to, user's acceptance of the system using standard usability tests, measurement methods need to be evaluated in order to derive and understand the effect of such framework on immersive reading experiences.

Figure 1.1: Multimedia Home Entertainment Scenario

The research in this thesis has been performed by leveraging several existing tech-

[1]Haptic sofa, http://d-box.com/en/home-theatre/d-box-ready-seating/

nologies. However, the idea of hapto-audio-visual augmentation, e-book based multimedia quiz annotation, necessary generic system architecture, semi-automatic annotation methodology, possible interaction methodology, prototype implementations, and detail experiments are original works. The contributions of this research provide an intuitive mechanism of integrating different useful modalities in eBook reading platform which is found to leverage immersive reading experience. Overall contributions can be summarized as the following:

1. Proposal of annotation based haptic-audio-visual feedback integration in traditional digital text contents in order to leverage total immersive experience.

2. Design and development of a generalized framework for such a Human-Computer-Interaction application that can annotate various multimedia data with eBook contents as well as interact with the ubiquitous environment.

3. Design and development of a multimedia ordering algorithm to meaningfully convey various media to the reader.

4. Design and development of multimedia quiz augmentation method in the proposed HE-Book system. Such quiz can be extended to any types such as Multiple Choice Questions, fill in the blanks, anagrams, crosswords, word searches etc.

5. Design and development of a semi automatic hapto-audio-visual annotation scheme by analyzing the text scenarios of the e-book content. For this purpose, we have developed noun phrase based video-audio search and rule based emotion detection methodology.

We also extended an open source based e-book reader to support HE-Book based extension and e-book authoring facilities. The annotation methodology used in the prototype allows the creation and storing of separate annotation file and corresponding multimedia resources in remote annotation repository. This will help to easily extend existing eBook platforms to our proposed HE-Book enabled platform. Finally, we present various analysis of the quantitative and qualitative test results carried using the prototype HE-Book systems.

1.4 Scholarly Output

In addition to meet its objectives as described above, this research undertaking has also lead to a variety of scholarly publications, as listed below.

1. **Papers in Referred Journals**

   (a) **Kazi Masudul Alam**, Abu Saleh Md Mahfujur Rahman, and Abdulmotaleb El Saddik, "Mobile Haptic Ebook System to Support 3D Immersive Reading in Ubiquitous Environments", Accepted (In Revision), ACM Transactions on Multimedia Computing, Communications and Applications, April, 2012.

   (b) Faisal Arafsha,**Kazi Masudul Alam**, Abdulmotaleb El Saddik, "Towards Consumer Centric Affective Haptic Clothes", To be submitted, Journal of Multimedia Tools and Applications, Springer, June, 2012.

2. **Papers in Referred Conferences**

   (a) **Kazi Masudul Alam**, Abu Saleh Md Mahfujur Rahman and Abdulmotaleb El Saddik, "HE-Book: A Prototype Haptic Interface for Immersive E-Book Reading Experience", IEEE - World Haptics Conference (WHC), pages 367-371, 2011, Istanbul.

   (b) Abu Saleh Md Mahfujur Rahman, **Kazi Masudul Alam** and Abdulmotaleb El Saddik, "Augmented HE-Book: A Multimedia Based Extension to Support Immersive Reading Experience", International Conference on Autonomous and Intelligent Systems (AIS), LNCS, pages 321-330, 2011, Burnaby, BC, Canada.

   (c) Abu Saleh Md Mahfujur Rahman, **Kazi Masudul Alam** and Abdulmotaleb El Saddik, "A prototype Haptic E-Book system to support immersive remote reading in a smart space", pages 124-128, Haptic Audio Visual Environments and Games (HAVE), 2011 IEEE International Workshop.

   (d) **Kazi Masudul Alam**, Abdulmotaleb El SADDIK,Sandro HARDY, Aysha AKTHER, "SMS Text Based Affective Haptic Application", Proceedings of Virtual Reality International Conference (VRIC 2011), 6-8 April 2011, Laval, France.

(e) **Kazi Masudul Alam**, Md. Abdur Rahman, Abdulmotaleb El Saddik, Wail Gueaieb, "Adding Emotional Tag to Augment Context-Awareness in Social Network Services", IEEE International Instrumentation and Measurement Technology Conference (I2MTC),pages 1-6, 2011

(f) Faisal Arafsha,**Kazi Masudul Alam**, Abdulmotaleb El Saddik, "EmoJacket: Consumer Centric Wearable Affective Jacket to Enhance Emotional Immersion", In Press, IEEE Innovations in Information Technology, 2012, March 18-20, UAE

(g) Aysha Akther, **Kazi Masudul Alam**, Heung-Nam Kim, Abdulmotaleb El Saddik, "Social Network Assisted Personalization with User Context for Recommender Systems", In Press, IEEE Innovations in Information Technology, 2012, March 18-20, UAE

1.5 Thesis Organization

The remainder of this thesis is organized as follows:

Chapter 2 presents an overview of the background literature and related studies. Background literature contains the importance of book in education, how e-book has emerged and taken the space of real book. Also, we describe the types of eBook readers available and the devices that support eBook platform. Next we discuss about multimedia augmentation and ubiquitous computing. Then, we present overview of the closely related research works that motivated our proposed system.

Chapter 3 presents our proposed system architecture and provides details about the component architecture and their interaction methods. Next we discuss about our semi-automatic eBook annotation scheme. It also includes emotion detection based recommendation phase which helps in the annotation process.

Chapter 4 gives a detailed description of the implementation process of this system. It explains the system setup upon which we have developed our prototype systems. Next we describe the software development details of the prototypes and the emotion detection model including some sample user interfaces. Finally, this chapter provides a section on the problems faced and lessons learned during the implementation phase.

Chapter 5 provides evaluation results and corresponding findings. We have divided the evaluation phase in qualitative and quantitative measurements to determine useful involvement methodologies with the users. Next we have conducted detail user study to

measure the influence of our system in immersive and effective reading experience.

Finally, in chapter 6 we summarize the overall contribution of our research and draw a path to possible future work to enhance the system and make it more intuitive to common user practice. We also discuss few research venues.

Chapter 2

Background and Related Works

2.1 Literature Review

In this section, we present a review of the literature related to E-Book system. Here we draw comparisons among existing eBook file formats and popular eBook hardware platforms. We also present short reviews about multimedia (text, image, sound, video, haptics) representation, immersion, multimedia authoring tools, multimedia augmentation and ubiquitous computing.

2.1.1 E-Book System

An **E**lectronic**-Book** (variously, e-book, ebook, digital book) is a book-length publication in the digital form, consisting of text, images, or both, and can be viewed on desktop computer, laptop, tablet, smartphone or e-book reader [21]. During travel, we can store a large number of e-books in portable units, which can dramatically reduce weight and volume compared to paper. As well as, electronic bookmarks make referencing easier, and e-book readers can allow the user to annotate pages.

The inventor and the title of the first e-book is not widely agreed upon. It is said that Thomas Aquinas created the landmark of the e-book in 1940 with his heavily-annotated indexing to the works. Alternatively, electronic books are considered by some to have started in the mid-1960s, with the NLS project headed by Doug Engelbart at Stanford Research Institute (SRI), and the Hypertext Editing System and FRESS projects headed by Andries van Dam at Brown University [59][60][4]. Yet others believe that the inventor of the e-book is Michael S. Hart through the Project Glutenberg in 1971 [31]. One early e-book implementation was the DynaBook e-book reader for notebook in 1970.

The major problem with e-books is the many formats competing for prime time, including Adobe PDF, iBooks, eReader, Mobipocket Reader, EPUB, Kindle and Microsoft Reader. However, in the late 1990s a consortium was formed to develop the Open eBook format as a way for authors and publishers to provide a single source document that could be handled by many book-reading software and hardware platforms. Table 2.1 shows a comparison[1] of features supported by some of the eBook file formats. Followings are the most widely used formats for eBook systems.

- ***HTML*** Hyper Text Markup Language (HTML) is the markup language used for most web pages. In this case, web browsers are used to read E-books written in HTML. World Wide Web Consortium (W3C) provides the specification of HTML format without any charge.

- ***Kindle*** In 2011, Amazon.com released Kindle Fire Reader. It included the next generation eBook file format Kindle Format 8. Both HTML5 and CSS3 are supported in *.kf8* file format. Amazons older Kindle eBook readers use proprietary format, *AZW*.

- ***Microsoft Reader*** Microsoft also has their own bites in the eBook market. Their proprietary Microsoft Reader Program supports Digital Rights Management (DRM) protected *LIT* files. LIT file format is somehow similar to CHM file format from Microsoft.

- ***EPUB*** International Digital Publishing Forum (IDPF) has created an open standard for the eBooks which is called *.epub*. *EPUB* file format has gained rapid popularity since its publication and by 2011 it has become the most widely supported vendor-independent XML based eBook format.

- ***Portable Document Format (PDF)*** Well known Adobe Systems also has their participation in the eBook space. Their *PDF* file format was first initiated to provide a standard form for storing printable documents, but now it is widely used as eBook file format. This format is derived from PostScript. It supports additional features like compression, passwords and Digital Rights Management (DRM) but not language features like loops.

[1]Wikipedia (Comparison of e-book formats), http://en.wikipedia.org/wiki/Comparison_of_e-book_formats

Format	Image Support	Sound Support	Open Standard	Embedded Annotation Support	Video Support	Haptics
HTML	Yes	Yes	Yes	No	Yes	No
Kindle	Yes	Yes	Yes	Yes	Yes	No
Microsoft Reader	Yes	No	No	Yes	No	No
EPUB	Yes	Yes	Yes	Yes	Yes	No
Portable Document Format (PDF)	Yes	Yes	Yes	Yes	Yes	No
eReader	Yes	No	No	Yes	Yes	No
Mobipocket	Yes	No	Yes	Yes	Yes	No

Table 2.1: Comparison of features supported by various eBook file formats

- ***eReader*** eReader is a freeware program for viewing Palm Digital Media electronic books which use the *pdb* format used by many Palm applications. Versions are available for iPhone, PalmOS (not webOS), Android, Symbian, BlackBerry, Windows Mobile Pocket PC/Smartphone, desktop Windows, and Macintosh.

- ***Mobipocket*** The Mobipocket e-book format is based on the Open eBook standard using XHTML and can include JavaScript and frames. It also supports native SQL queries to be used with embedded databases.

2.1.2 Popular E-Book Platforms

Many millions of eBook readers are using the following list of popular eBook platforms (Figure 2.1) in their everyday reading experience. These devices are widely accepted in the existing readers' community and table 2.2 shows the range of file formats supported by these ebook reading devices.

- ***Amazon Kindle Fire*** The Kindle Fire[2] is a tablet computer version of Amazon.com's Kindle e-book reader. It was announced on 28 September 2011, which

[2]Wikipedia (Kindle Fire), http://en.wikipedia.org/wiki/Kindle_Fire

has a color 7" multi-touch display. It supports IPS technology and uses a forked version of Google Android Operating System.

Reader	**PDF**	**EPUB**	**HTML**	**MobiPocket**	**eReader**	**Kindle**	**Open eBook**
Amazon Kindle Fire	Yes	Yes	Yes	Yes	No	Yes	No
Android eBook Reader	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Apple eBook Reader	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Barnes & Noble Nook	Yes	Yes	No	Yes	Yes	No	No
Kobo eReader	Yes	Yes	Yes	No	No	No	No
Sony Reader	Yes	Yes	No	No	No	No	No
Windows Phone 7	Yes	Yes	Yes	No	No	Yes	No

Table 2.2: Comparison of different file formats supported by various eBook platforms

- ***Android eBook Reader*** Android [3] is a Linux-based operating system for mobile devices such as smartphones and tablet computers. It is developed by the Open Handset Alliance led by Google. Google purchased the initial developer of the software, Android Inc., in 2005. It has multiple ebook reading software available in its market place.

- ***Apple eBook Reader*** iBooks[4] is an e-book application by Apple Inc. It was announced in conjunction with the iPad in January, 2010, and was released for the

[3]Wikipedia (Android), http://en.wikipedia.org/wiki/Android_(operating_system)
[4] Wikipedia (iBooks Author), http://en.wikipedia.org/wiki/IBooks_Author

iPhone and iPod Touch in mid-2010, as part of the iOS 4 update. It primarily receives ePub content from the iBookstore, but users can also add their own ePub and PDF files viadata synchronization with iTunes.

Figure 2.1: Popular Table / E-Book Reader Devices available in the market currently.

- ***Nook*** The Barnes & Noble Nook[5] is a brand of electronic-book reader developed by American book retailer Barnes & Noble, based on the Android platform.

- ***Kobo*** The Kobo eReader[6] is an e-book reader produced by Toronto-based company Kobo Inc. The original version was released in July 2010 and was marketed

[5]http://en.wikipedia.org/wiki/Barnes_%26_Noble_Nook
[6]http://en.wikipedia.org/wiki/Kobo_eReader

as a minimalist alternative to the more expensive e-book readers available at the time.

- ***Sony Reader*** The Sony Reader[7] is a line of e-book readers manufactured by Sony. It uses an electronic paper display developed by E-Ink Corporation, is viewable in direct sunlight, requires no power to maintain a static image, and is usable in portrait or landscape orientation.

2.1.3 Multimedia Representation

Multimedia is a woven combination of digitally manipulated text, photographs, graphic art, sound, animation and video elements [66]. Followings are some types of media that can be compiled together to form multimedia presentations.

- ***Text*** Words and symbols in any form, spoken or written, are the most common means of communication. Text is an element of multimedia menus, navigation systems, and content.

- ***Image*** Still image is another important element of a multimedia project. While creating still images, display resolution, hardware and software capabilities are considered. At the same time, still images can be generated in two ways: bitmaps, vector-drawn graphics.

- ***Video*** Light reflected from an object through the camera's lens is converted into the electronic signal by a charge-coupled device. This electronic signal contains three channels of color information and synchronization pulses. Several video stands exist that deal with the amount of separations between the components of the signal. Video is a popular tool for delivering multimedia.

- ***Sound***Vibrations in the air create waves of pressure that are perceived as sound. Sound waves vary in sound pressure level (amplitude) and in frequency or pitch. For multimedia systems, sounds are assigned to various system events such as startup, warnings, etc.

- ***Haptics*** Haptics, a term that was derived from the Greek word haptesthai meaning of a relating to the sense of touch, refers to the science of manual sensing

[7]http://en.wikipedia.org/wiki/Sony_Reader

(exploration of information extraction) and manipulation (for modifying the environment) through touch. Haptic technology is an emerging interdiscipline field that deals with the understanding of human touch (human haptics), motor characteristics (machine haptics), and with the development of computer-controlled systems (computer haptics) that allow physical interactions with real or virtual environments through touch [19].

2.1.4 Multimedia Authoring

Multimedia authoring tools enable the developer to create, edit, and import data. In multimedia authoring systems, multimedia elements and events are often regarded as objects. Objects exist in a hierarchical order of parent and child relationships. Each object is assigned properties and modifiers. Modifiers such as methods, functions, etc mini-programs are used to modify the properties of the object. On receiving messages, objects perform tasks depending on the properties and modifiers. Authoring systems include editing tools to create, edit, and convert multimedia elements such as animation, audio and video clips. The organization, design, and production process for multimedia involve storyboarding and flowcharting. Visual programming with icons or objects is the simplest and easiest authoring process. Authoring tools offer very high level language (VHLL) or interpreted scripting environment [66].

2.1.5 Immersion

Immersion[8] is the state of consciousness where an immersant's awareness of physical self is diminished or lost by being surrounded in an engrossing total environment. This mental state is frequently accompanied with spatial excess, intense focus, a distorted sense of time, and effortless action. The term is widely used for describing immersive virtual reality, installation art and video games. According to Ernest W. Adams immersion[9] can be separated into three main categories:

- *Tactical immersion* Tactical immersion is experienced when performing tactile operations that involve skill. Players feel "in the zone" while perfecting actions that result in success.

[8]Immersion (virtual reality), http://en.wikipedia.org/wiki/Immersion_(virtual_reality)

[9]Postmodernism and the Three Types of Immersion ,http://designersnotebook.com/Columns/063_Postmodernism/06

- *Strategic immersion* Strategic immersion is more cerebral, and is associated with mental challenge. Chess players experience strategic immersion when choosing a correct solution among a broad array of possibilities.

- *Narrative immersion* Narrative immersion occurs when players become invested in a story, and is similar to what is experienced while reading a book or watching a movie.

2.1.6 Multimedia Augmentation

Augmented reality (AR) [10] is a live, direct or indirect, view of a physical, real-world environment whose elements are augmented by computer-generated sensory input such as sound, video, graphics or GPS data. It is related to a more general concept called mediated reality, in which a view of reality is modified (possibly even diminished rather than augmented) by a computer. Augmented Reality is considered as an extension of Virtual Reality. Virtual Reality (VR) is a virtual space where the player immerses themselves into that exceed the bounds of physical reality (Figure 2.2). In the VR, time, physical laws and material properties no longer hold in contrast to real-world environment.

Augmented Reality might apply to all senses, not just sight. AR could be extended to include sound. The user would wear headphones equipped with microphones on the outside. The headphones would add synthetic, directional 3D sound, while the external microphones would detect incoming sounds from the environment. This would give the system a chance to mask or cover up selected real sounds from the environment by generating a masking signal that exactly canceled the incoming real sound.

Another example is haptics. Gloves with devices that provide tactile feedback might augment real forces in the environment. For example, a user might run his hand over the surface of a real desk. Simulating such a hard surface virtually is fairly difficult, but it is easy to do in reality. Then the tactile effectors in the glove can augment the feel of the desk, perhaps making it feel rough in certain spots. This capability might be useful in some applications, such as providing an additional cue that a virtual object is at a particular location on a real desk [6].

[10] http://en.wikipedia.org/wiki/Augmented_reality

Figure 2.2: A sample augmented reality scene, where 3D representation of a 2D player is visible through some head mounted / hand held device.

2.1.7 Ubiquitous Computing

Ubiquitous computing[11] is a model of human-computer interaction in which information processing has been thoroughly integrated into everyday objects and activities. In the course of ordinary activities, someone "using" ubiquitous computing engages many computational devices and systems simultaneously, and may not necessarily even be aware that they are doing so. This model (Figure 2.3) is usually considered an advancement from the desktop paradigm. More formally, Ubiquitous computing is defined as "machines that fit the human environment instead of forcing humans to enter theirs". This paradigm is also described as pervasive computing, ambient intelligence, where each term emphasizes slightly different aspects. When primarily concerning the objects involved, it is also physical computing, the Internet of Things, haptic computing, and things that think [52].

[11] http://en.wikipedia.org/wiki/Ubiquitous_computing

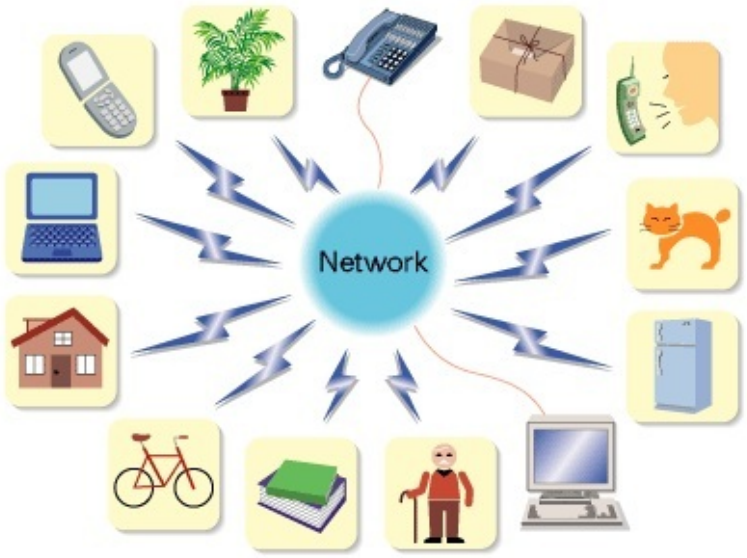

Figure 2.3: Ubiquitous computing will enable diverse wireless applications, including monitoring of pets and houseplants, operation of appliances, keeping track of books and bicycles, and much more.

2.2 Related Works

According to a research article by Harrison [25], French visionaries Robida and Uzanne stated the future of books in 1895. Their prognosis was that written book contents would soon be transferred into sound wave and be played in gramophone and subsequently gramophone would evolve to become a pocket-sized player. They also envisioned that the recorded voice of the books will be maintained by one authority and rented by restaurants, public transportation vehicle, steamship cabin, hotel room, etc. Though the future of reading didn't directly evolve like they thought but certainly we found various similarities with actual book reading trends.The first portable, wireless, networked device that could act as a notebook/reading device was introduced by Alan Kay in 1968. This is called DynaBook (Fig. 2.4) and still considered as an ultimate reading device. Though the model was never physically built, but it did set the stage for the ebook devices. In the mid-1990s, couple of book-like metaphors emerged such as SoftBook Reader (Fig. 2.5) , Rocket eBook (Fig. 2.6) etc. Their features included the latest memory, processor, and LCD technology as well as network connection, monochrome/color display etc. Considerable design efforts were made to give them book-like reading experience. Hence, we find that several research efforts have been put in this research area for long time of

computing history, which acted as guidelines for our research work. In this section, we briefly describe some closely related systems, their basic concepts as well as following architecture.

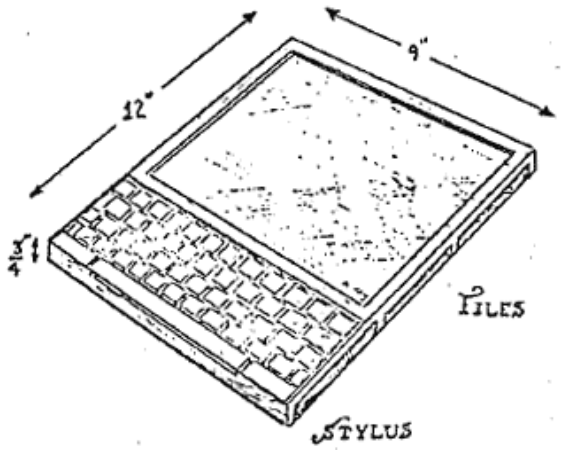

Figure 2.4: Dynabook is the first proposed eBook device. [25]

Figure 2.5: SoftBook eBook reader device. [25]

2.2.1 MagicBook

In contemporary research, MagicBook [8] project was an early (2001) attempt to bring augmented reality in real book reading experience, where a user could see 3D virtual

models out of the real book pages. If a person looks at the pages through an augmented reality display, they see 3D virtual models appearing out of the pages (Figure 2.7b). Users can move themselves or the book itself to go through the augmented reality scene. So basically, it is an enhanced version of the traditional 3D pop-up book.

Figure 2.6: Rocket ebook reader device. [25]

In the MagicBook, users can change the virtual models as well, just by turning the book pages. If someone likes one specific page, he/she can get into the virtual scene, move around and can have immersive experience into a story. Another interesting advantage of the system is that, multiple users can immerse in the same scene where one user visualizes others as miniature avatars (Figure 2.7c). In the immersive world, people viewing the augmented reality scene appear as large, virtual heads looking down from the sky. This way, people are always aware of the other users of the interface and where they are looking.

The MagicBook interface uses a handheld augmented reality display (HHD), a computer graphics workstation, and the physical book. Each user has his/her own handheld display and computer to generate personal scene views. For multiple user augmentation, these computers need to be networked together to exchange avatar related information. These computers are networked together to exchange information about avatar positions and the virtual scenes each user view [8].

2.2.2 3Book

Card et al.(2004) [11] proposed physical behavioral simulation of a real codex book in the presentation of a 3D virtual book which is called 3Book. Later, Welch et al.

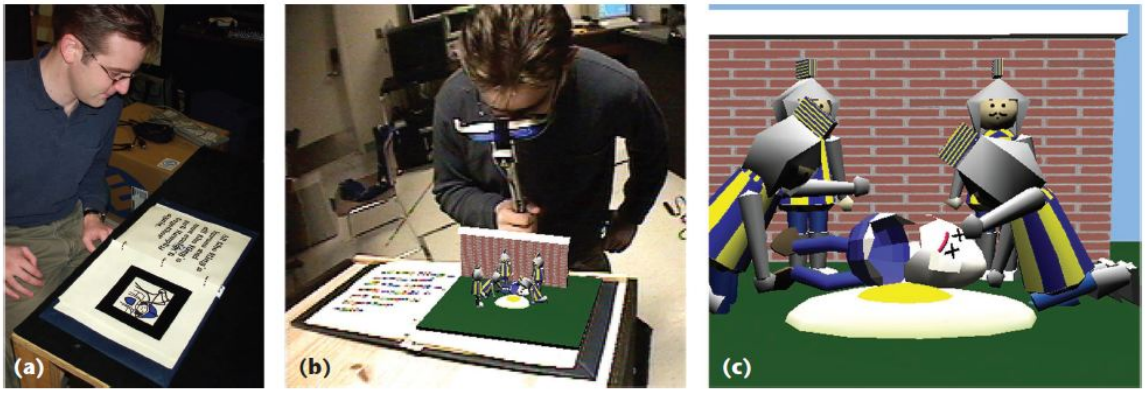

Figure 2.7: Using the MagicBook interface to move between reality and virtual reality. (a) Reality, (b) augmented reality, and (c) immersive VR. [8]

(2005) [68] extended the concept further to an immersive Virtual Reality book aimed at medical training for surgery. 3Book can be used in the digital library and several other information-intensive applications. It could represent books of unlimited length, supported both short and large format of books as well as helped in reading enhancement and sense making.

Authors started their virtual book work with the system Catalogues [12], which was used for TabWorks [39]. Later they used WebBooks [13], a more elaborated book metaphor where web pages were piled in a 3D workspace. Analyzing the above systems, they found two limitations; first, they were not supporting interactive 3D virtual book and second, they could only support limited number of pages, as well as they were relatively slow. Hence they developed their 3Book prototype, which simulates physical codex book metaphor in 3D electronic format. First, they developed their system in a paper book named *Biohazard* which had 300 pages (about 118,000 words). The reader could turn individual pages by touching them or turn blocks of pages by touching the core edge of the book. They used VTK 4.2 and JAVA programming language for these purposes.

Their system also supported full size larger books such as *Readings in Information-Visualization: Using Vision to Think* of pages (700), a large page format (8-1/2 x 11) containing 700 1100 words per page, for an estimated 700,000 words. It was difficult for a large laptop with an LCD screen of 9" by 12" to view Information Visualization, which required them to reduce the size by 80%. To resolve these problems, their system supported rapid zooming of the book and display the pages in rocker page form. If there

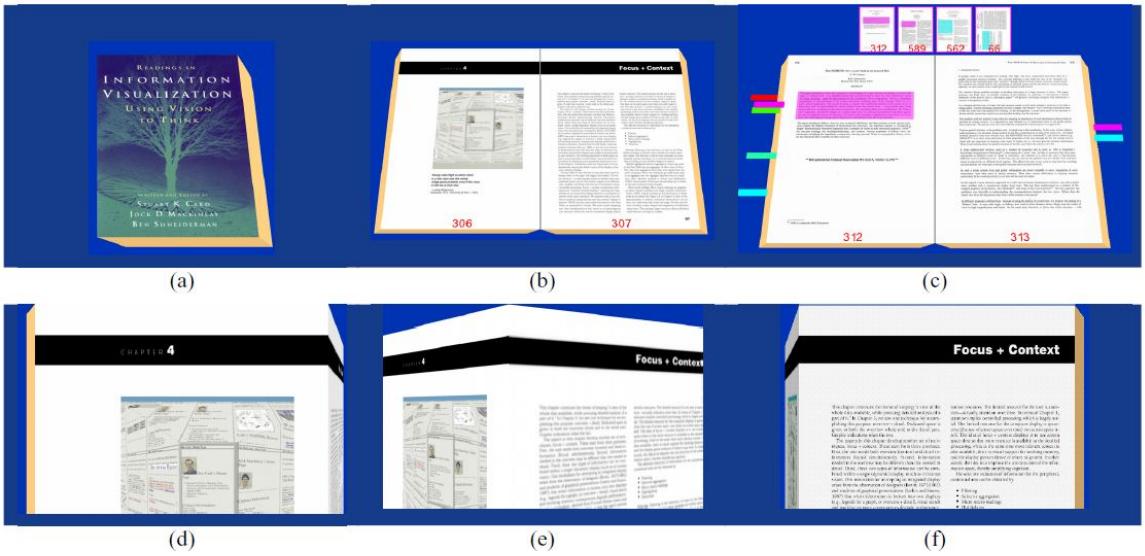

Figure 2.8: 3Book rocker pages. (a) Large 3Book, (b) In open position, (c) 3Book with highlights, bookmarks, and slid-out pages, (d) Recto page bent back and verso page zoomed into reading position, (e) Rocker page in animated transition from verso to recto page, (f) Recto page in reading position. [11]

is not enough horizontal room, the non-viewed page is bent back, allowing all of its page width still to be visible. 3Book also supported bookmarking and sensemaking activities, which allowed the comparison of the multiple pages. They provided smart indexes based on information scent and word co-occurrence models that dynamically changed in response to reader interest.

2.2.3 Annotating 3D Electronic Books

The same researcher group of 3Book, extended their system in [27] (2005), to support 3D electronic book annotation. In that system, they adopted two types of annotation: *highlight* and *scribble*. A *highlight* is depicted as an area on the page, which can be used to highlight a figure, a paragraph, or even a previously added annotation in the margin. On the other hand, a *scribble* is presented as an annotation made with pen strokes. Examples of scribble are underlined words and sentences, brief notes written between lines, and marginal symbols like arrows and brackets. In order to render the annotation, they used three different techniques.

- *Transparent geometry* Transparent geometry (e.g., polygons or lines) are repre-

sented as annotation on the pages.

- *Vertex coloring* The color values at the vertices of the polygonal mesh of page are modified for this annotation.
- *Texture coloring* Pixels of a page texture is updated to achieve this effect.

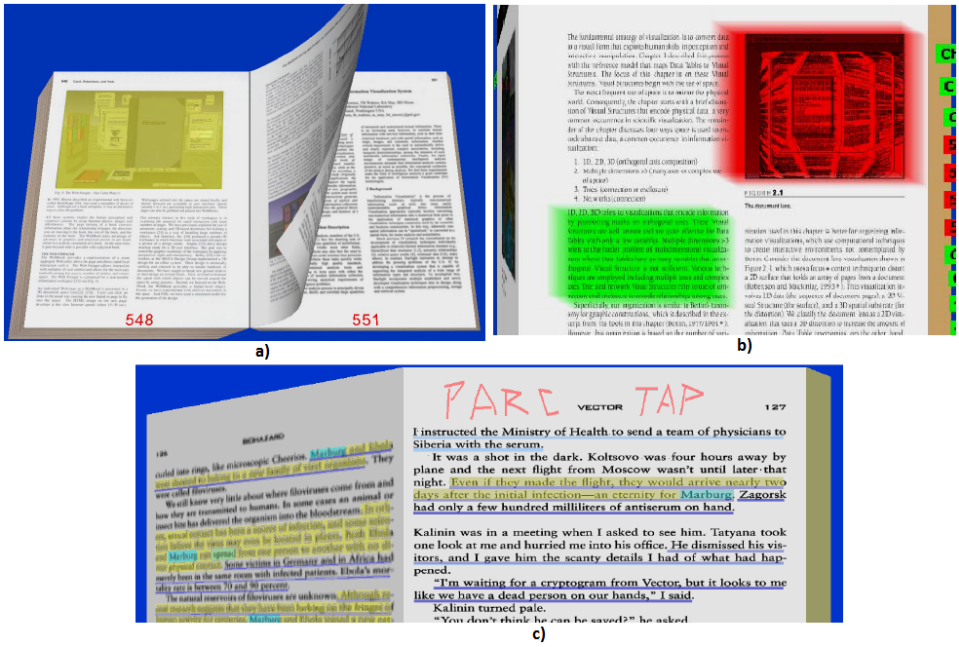

Figure 2.9: (a)A yellow, transparent 3D rectangle is placed on top of the left page. A transparent geometry would have to be deformed along with the page area that it annotates, (b) Two page areas are highlighted in red and green, respectively, (c) 3Using the texture coloring technique, PARC TAP was scribbled in the top margin of page 127. Note how system response time limits quality of annotation with this technique. [27]

For their system, authors combined the above techniques in several ways to improve the rendering speed at different cases. For example, when a user is sketching an annotation, their system shows a 3D *transparent geometry* instead of updating the page and later map the annotation onto the texture. On the other hand, for stroke based annotation, when user is drawing the trajectory using the pen, they create polyline in 3D scene. When an user releases the pen, then they update the page texture using coloring technique. The goal of this work is to investigate suitable graphical and user interface techniques for supporting interactive annotation in 3Book.

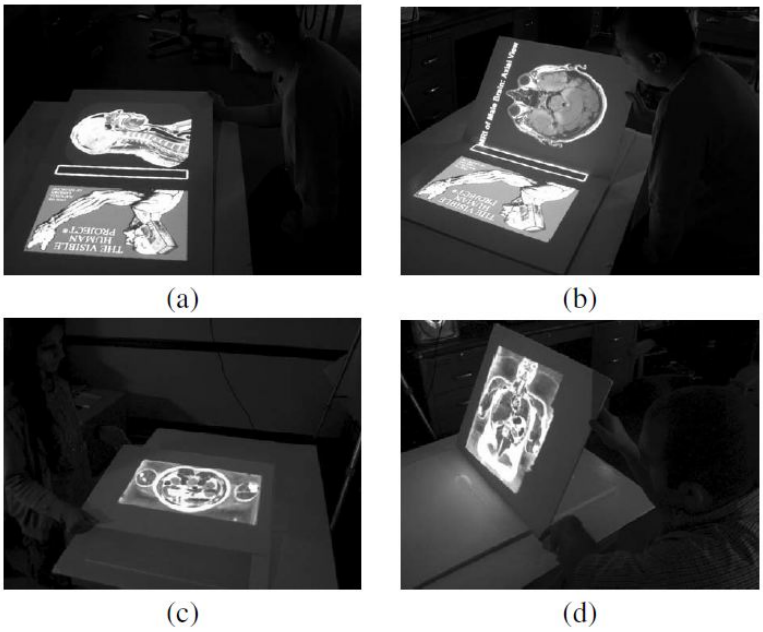

Figure 2.10: A projector augments the Universal media book with data from the Visible Human Project. (a) A user looks at a sagittal view image of a face. (b) A video fly-through of axial mri scans of a skull is played and remains correctly aligned to the page extents while a user turns the page. (c), (d) A user dynamically explores an anatomical volumetric dataset of a male cadaver through varying directions and orientations.[23]

2.2.4 The Universal Media Book

In another research, Gupta and Jaynes (2006) [23] explored integration of projected imagery with the physical book in order to introduce a tangible interface to multimedia data. They have used camera and projector pair to introduce a tracking framework, which monitors the 3D position of planar pages when a user turns book pages back and forth. Rectangular white boards are stacked and bounded on one side to make up the book. Four 3D coordinates defined each of the book pages. Two of the coordinates, called pivot points, remain fixed at the spine, while the other two corresponding coordinates, called endpoints, are constantly updated as a page is turned. As pages are turned, they are assumed to follow a path along the curved edge of a cylinder whose major axis is parallel to the book spine and is embedded in the tabletop (Figure 2.10). As a page is moved, the pages location in the world must be continuously updated in each frame so the projector can correctly augment the page. Their book can be loaded with various multimedia contents like video, images, data, etc.

2.2.5 Virtual Pop-Up Book

Nobuko et al. (2007) introduced the Virtual pop-up book [62] system which is based on augmented reality technology. It is capable of displaying 3D virtual objects on a real book based on pose and position estimation of a camera. In their system, they do not use any markers because a picture book with markers looks unattractive. Their proposed system has several advantages namely; 3D rendering helps readers understand scenes. Secondly, characters look lively by using motions. Thirdly, author of the picture book can use the new representation, which mixes 2D and 3D rendering. Lastly, it can express time changes using animation.

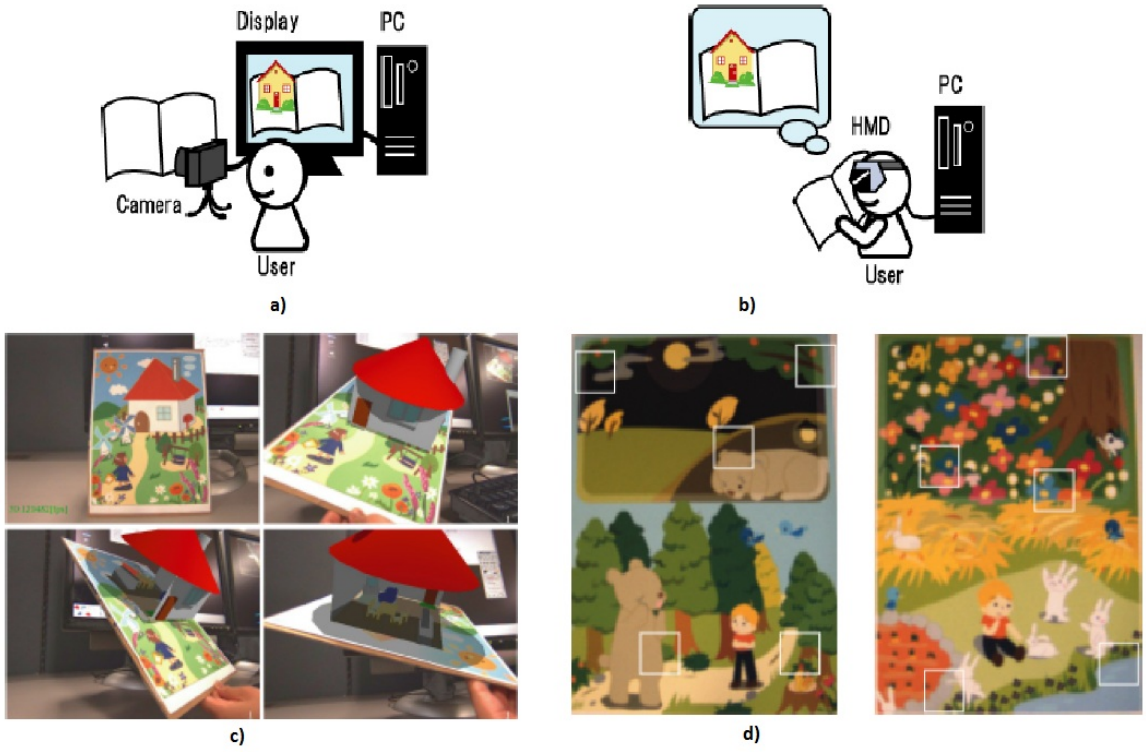

Figure 2.11: a)PC monitor-based display system , b)HMD-based display system, c)Viewing a scene from different angles, d)Example of selected templates. [62]

Their system had two kinds of display types. One is PC monitor (Figure 2.11a). The other is Head Mounted Display (Figure 2.11b). Their PC monitor version of the system is better suited for the children who feel HMD as a load. On the other hand, HMD version of the system offers more immersive experience to one user.

They used ARToolKit [29] for obtaining an image from the camera and rendering

Computer Graphics (CG) objects. It is very important to get the camera pose and position to the book, for rendering 3D CG objects at the proper position. They have used marker-less registration method in the contrary to many vision-based methods proposed until then. Their system flow has three parts Page determination, Initial pose estimation and Tracking.

- ***Page Determination*** For this phase, they create an image template of each page. The templates were cut out from input image by using image-editing software, later reduced 1/16 in size, searched using active search and finally, the highest similar template was adapted as the target page.

- ***Initial Pose Estimation*** The flow of the initial pose estimation follows: making templates of feature points (Off-line processing), searching some feature points from the input image, calculating camera position by using image coordinates of detected feature points.

- ***Tracking*** They applied texture tracking for the tracking method. This method does not use special markers in the iterative tracking phase. Several natural feature points were extracted from the tracked object by off-line image analysis. This phase follows the flow: forecasting the image to coordinate of the feature points, making templates based on the image coordinate of the feature points, template matching using the generated templates. After this procedure, they find the pairs of the world coordinates and image coordinates. With over three pairs, the camera poses can be acquired.

2.2.6 Dual-Display E-Book Readers

In order to replicate physical metaphors used in dealings with real books, Chen et al. (2008) [15] introduced a new design of dual display for e-book reader and various supported interactions like folding, flipping, and fanning in their system. They explored the design interfaces that support reading activities with the dual-display e-book reader. For this purpose, they build a prototype dual-display e-book reader integrated with motion sensors. Their device explored two unique aspects.

- ***Navigation using an embodied interface*** Their device provided similar set of physical interactions that users normally exhibit with a real book. Their device allowed readers to turn pages by fanning one face toward the other and, in the back-to-back configuration, by flipping the device over.

- ***Flexible display configurations*** Authors figured out the drawbacks of the existing e-book readers, and the summarized that they offered very limited screen space. Though it is possible to construct devices with bigger displays, but the increased size introduces cumbersome to use and carry around. Their proposed Dual-display e-book reader offered a compromise between increase screen and flexible, convenient form factor. When, two of their displays were attached, they increase the screen space Fig. 2.12 and represented a broader view of the document.

Figure 2.12: Prototype e-book reader with faces in the attached, side-by-side configuration (top), and the detached configuration (bottom). [15]

The main goal of their work was to explore, how dual-display can improve reading experience for digital content. They established three design goals based on ethnographic work on reading.

- ***Improving local navigation*** While reading, users need to consult on or around the current page repeatedly. Dual-display e-book readers can increase the reading area and provide lightweight, gesture based control which in turn facilitate in the reading process.

- ***Improving global navigation*** Sometimes book readers need searching and direct moving facility in the reading document. Dual-display can utilize its extra space for this purpose, which will improve the global navigation system.

- ***Improving multi-document navigation*** Many reading activities require simultaneous interaction with multiple distributed pages of a document. Furthermore, reader needs to take notes, references with the real book. Dual-display systems with detachable faces can improve multi-document navigation by allowing users to arrange and interact with each display separately.

2.2.7 Authoring Tool for Digilog Book

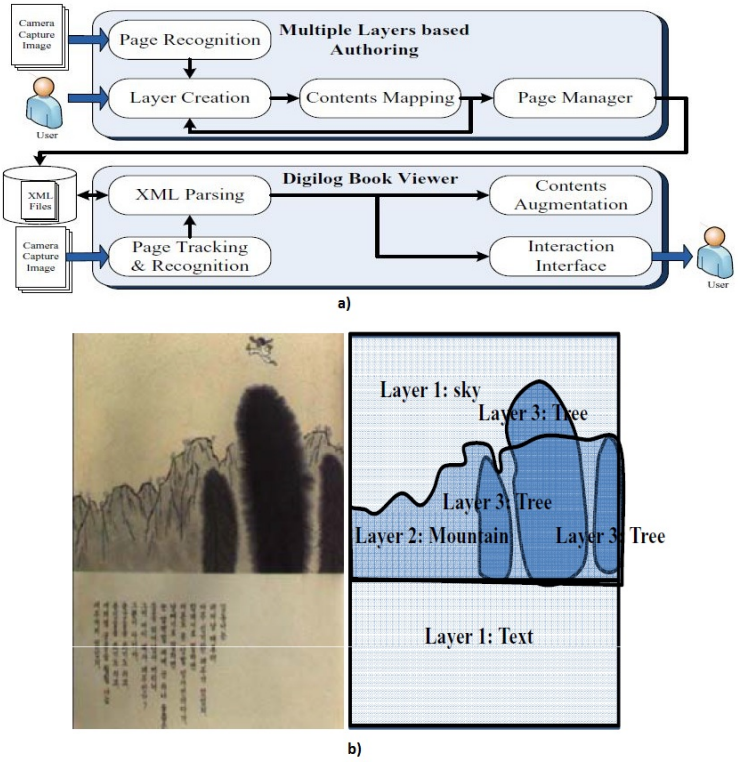

Figure 2.13: a)Overall Architecture, b) Description type of multi-layer of picture areas. [45]

Digilog Book (2007) [32] is an Augmented Reality (AR) book, which provides additional information by stimulating humans five senses with multimedia contents. In this system, user can manipulate virtual objects by various tangible pads. Park et al.

(2009) [45], proposed a multi-layer based authoring tool for Digilog Book (Figure 2.13). Their major feature was authoring various properties of printed contents of a paper book. Their architecture consists of two components, first a multi-layer (video, sound, picture, text) based authoring system and the other one is the DigiLog viewer. Their authoring phase starts with recognizing a page number from camera input image, then creating layers based on layer templates and finally storing the information using XML format. User can interact with the system when a page containing XML data is retrieved from the database.

2.2.8 Guidelines for Future Augmented Books

In a long-term development phase, Grasset et al. (2008) [22] have focused on design and development of mixed reality books in a broader sense. They have discussed the semantics of mixed reality book, available design space and user experience with such type of interfaces. They have found through study that contents typically found in augmented books are 2D static (images, schemes, text), 2D dynamic content (videos, animations), 3D content (environment, animation, 3D avatars), and sound (ambient, spatial, interactive). According to the authors, physicality of a book can be determined by how many physical elements are present. A *Virtual Book* is completely electronic format and *ARBooks* use physical book as an interface. In *Virtual AR Books* we can present virtual content only, multimedia side by side or multimedia integrated with the real book. Augmented books can support various tracking choices such as flat-rigid surfaces, curvature at the crease of the pages as well as non-rigid surfaces. In a mixed reality book, layout of virtual objects (2D, 3D, or 3D sound) can be of different geometric characteristics, which can enhance, replicate or replace 2D real book objects. We can interact with the virtual or mixed reality books by gaze, finger or tangible (positioning or moving tangible element) interactions. In order to provide immersive reading experience we need to consider invisible tracking, introduce visual surroundings and appealing cinematographic effects.

Park et al. (2010) [44] suggested a framework of book evolution for ubiquitous computing age. They proposed three dimensions (Figure 2.14) of book evolution : digitalization, augmentation, and hypermediation. The mainstream dimension of book evolution has been digitalization; that is, evolution of paper books to e-book.The recent development in e-book, is the emergence of e-book hardware and services integrated as a product service system. Like the product service system, i-tunes.com coupled with iPod of Apple

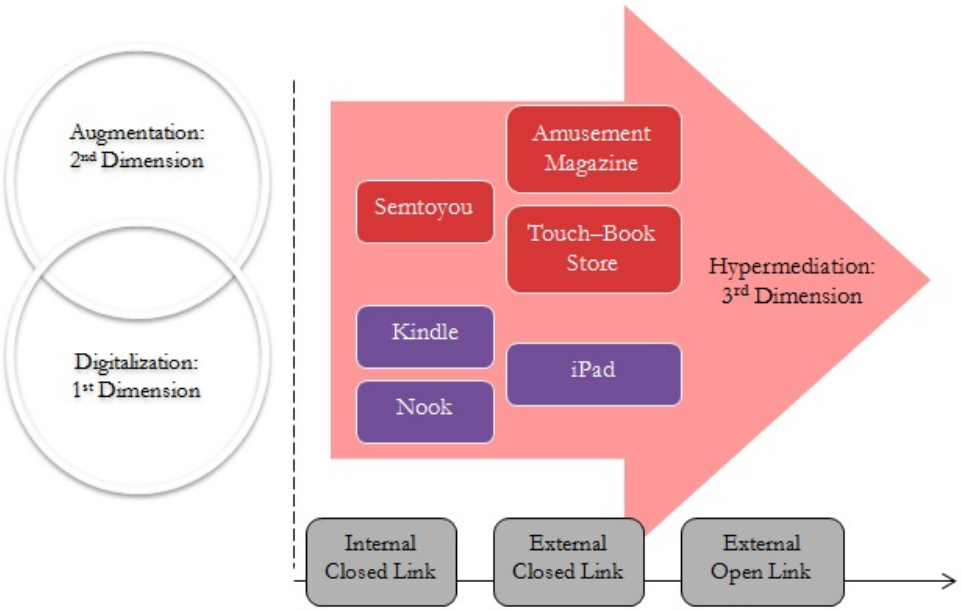

Figure 2.14: The Three Dimensions and Examples in Industry [44]

in music industry, e-book industries are competing with Amazon.com's Kindle, Barnes & Nobles Nook, Apples iPad with iBook. The second dimension of book evolution is augmentation: the evolution of paper books to augmented books. While the current e-book is digitalizing the current two-dimensional printed book into the two-dimensional computer screen, the augmented book is augmenting a different book experience by adding some dimensions or functionality to the traditional book or electronic book. The 3rd dimension is hypermediation where books are augmented with media using the standardized communication protocol such as HTTP.

2.2.9 Emotion Classification Methods

Researchers differ in the method of emotion classification but most of the theorists normally address love, joy, surprise, anger, sadness, and fear as basic emotions [64][51][28][43]. From Parrots [47] research article, we find that emotion can be presented in a tree structure where emotions are listed as primary, secondary and tertiary emotions. According to Parrot joy can be of type cheerfulness, pride, relief etc. and love can be of type affection, lust, longing. If we analyze the emotional labels of the following three sentences, Take my love, Love you :-X honey!(:-X denotes a big wet kiss), Loveeeee to see u soooon! using the methods provided in [28] and [42] we will label them all with love, however

if we use Parrot model [47], we can conclude that the above sentences intend affection, lust and longing respectively. Hence, this model is a good candidate for fine grained detail emotional tagging which can be used for a wide range of application domains e.g. handheld small device, serious games, haptic interaction, detail 3D emotion modeling, complex robotics etc.

A very important work for emotion analysis is affective labeling of WordNet [37] synsets which is called WordNet-Affect [56]. In WordNet-Affect noun, adjective, verb and adverb of WordNet are labeled to emotion, cognitive state, trait, behavior, attitude, feeling etc. WordNet-Affect follows the emotion model of Ortony and Turner [43]. Though Chuang [73] proposed a semantic network based text to emotion extraction method but it lacks necessary corpus for better results. Mulder et al. [41] made an attempt to formalize the affection in word and grammatical level of a fraction of Dutch and English. They tested their system in a pilot experiment but they do not provide their detail formalization or implementation methods.

A very popular work about affective haptic is iFeel_IM! [65] where 9 basic emotions are detected from SecondLife[12] chatting text and the emotion is conveyed to the end users using a set of devices. This research work is limited to the domain of SecondLife and their emotion detection does not support detail emotion tracking. HugMe system [18] facilitates users hugging using a specific haptic jacket while they are using instant messenger. This work does not have automatic emotion awareness. We have used this haptic jacket in our prototype application. Rehman and Liu conducted research [54] to support visually impaired people by wearing video camera that points to the face of the person of interest. Then the captured video is analyzed and facial emotional parameters are extracted and this information is conveyed to a mobile phone to vibrate based on that information.

2.3 Summary

We find from [44] that, several works on augmented books have explored the relationship between real and digital books, their interaction techniques and technologies and audio-visual augmentation. But the limitation of most of the existing researches on augmented books, lies in the requirements of specific and unpopular computer interface such as head mounted display, ARToolkit, bar code scanner, projector with cameras, etc. From

[12]Second Life, http://www.secondlife.com

the literature review, it is clear that there is hardly any existing research which has explored the possibilities of augmenting haptics with eBook reading materials intuitively and explained the effect of such modality in immersive reading experience. As a result, there is no such eBook reading platform which leverages the advantage of a ubiquitous smart space where there is haptic devices (i.e. jacket, sofa, arm band), multimedia entertainment systems (i.e. TV, Sound system). Hence, in this thesis, we have proposed our HE-Book system, which can fill that gap.

Chapter 3

HE-Book System Design

This chapter presents the overall architecture and detailed descriptions of a new annotation based e-book reading system, which supports various types of augmented traditional multimedia (image, video, audio), haptic, quiz and can easily adapt to new media. When a user reads an e-book in desktop, mobile, or tablet device using the proposed Haptic E-Book (HE-Book) reader, he/she will perceive different types of modality, which help to enhance reader's immersive reading experience. Such total(visual, audio, haptic) immersion helps one reader to properly understand and retain the content of e-book materials. We have illustrated our system description in the section 3.1 and provided e-book annotation scheme in section 3.2.

3.1 Haptic E-Book System

The HE-Book system consists of three major modules and an interaction controller as depicted in Figure 3.1. All the modules are composed of different components, which is pointed in Figure 3.2. In Section 3.1.1, we describe *HE-Book Event Manager* module, which assists in reading event detection, locating target paragraphs based on the user's finger touch or automatic reading interaction. It also describes the details about runtime quiz display and following answer event management. In Section 3.1.2, we describe part of the *Annotation Manager* module that handles e-book annotation retrieval operations. *Multimedia and Haptic Rendering Manager* module, illustrated in Section 3.1.3, describes details about diverse multimedia, haptic and quiz rendering procedures. Operations of all the modules, that is the interaction among various components are centrally monitored and organized by the *Interaction Controller*. It is described in Section 3.1.4.

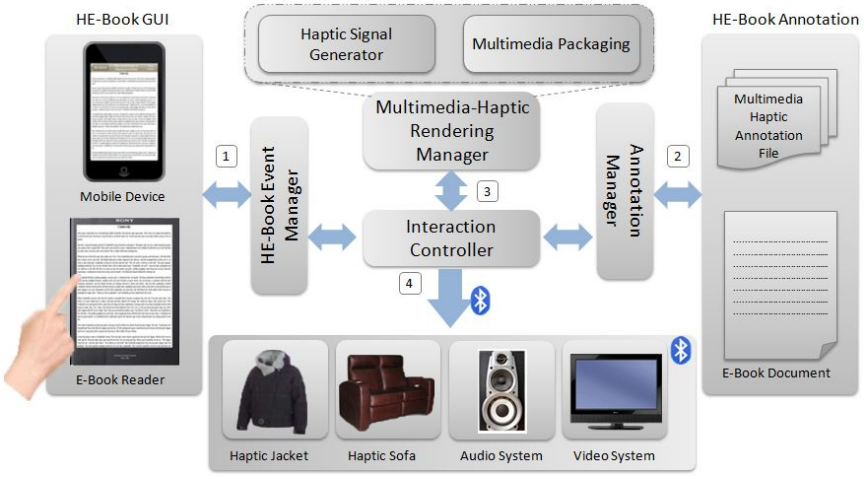

Figure 3.1: High level architecture of the proposed HE-Book system

3.1.1 HE-Book Event Manager

In this section, we describe reading event generatation procedure by user interaction and their corresponding processing methodologies. In HE-Book, each reading event request can initiate multiple event processing responses. Reading event can be an outcome of user‘s finger touch or machine automated reading itself. For each reading event there can be diverse multimedia presentation as well as a series of quiz interaction.

Reading Event Generation Methodologies

- *Finger Touch Based:* The haptic e-book system can be deployed in a touch based desktop/mobile/tablet device. The user can flip pages and scroll the paragraph of the pages through the touch based interaction. If the HE-Book reader system is used in a touch mobile device (Figure 4.5), the standard touch SDK available for that device is used to obtain the screen coordinates, using the *Reading Event Listener* component. The Listener sends the corresponding device specific touch coordinate information to the *Reading Event Manager Interface* (Figure 3.3). This component uses the *Paragraph Locator* component in order to match the touched screen coordinate with the displayed e-book page area to deduce the paragraph that the user is currently pointing at. As soon as the user touches a paragraph of the page that has been previously annotated, the user is sent a motor vibration to confirm that a valid haptic annotation for the selected paragraph has been found. Further processing is controlled by the*Interaction Controller* module.

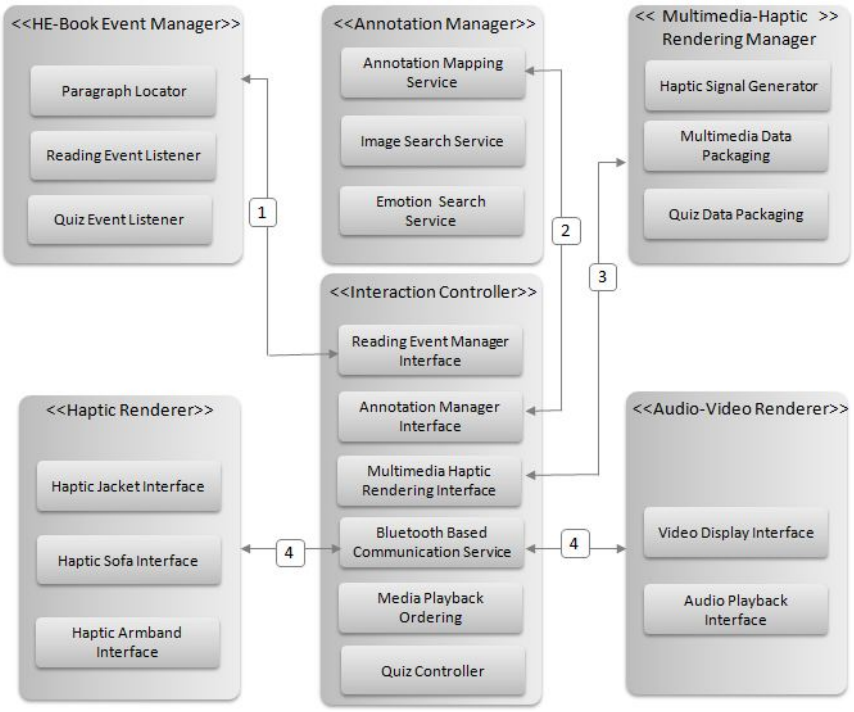

Figure 3.2: HE-Book component architecture

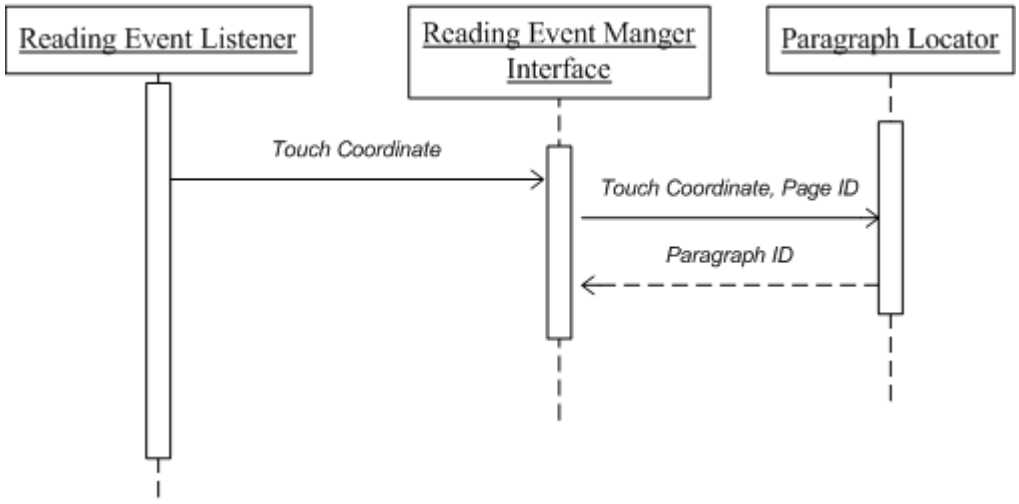

Figure 3.3: Sequence diagram of HE-Book reading event

- *Machine Initated Reading:* With the vision of a user friendly system for adult or children, physically well or challenged, we incorporate machine initiated reading support in the system. We consider machine initiated reading settings, is useful for

the young children or physically challenged people. Using this automated interaction mode, a user can choose to listen selected or all the remaining paragraphs of the book automatically. Reading event is started from the current paragraph from which Text-To-Speech (TTS) system is activated. When the automatic reader system selects an already tagged paragraph, the *Paragraph Locator* component comes to know about the corresponding paragraph for which multimedia-haptic augmentation needs to be played. Next, *Interaction Controller* module is handed over to the control to process rest of the operations.

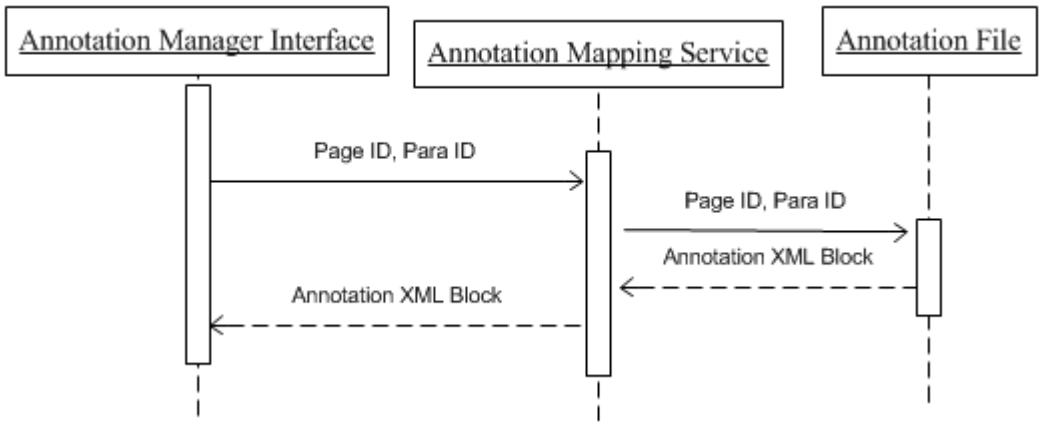

Figure 3.4: HE-Book annotation mapping sequence diagram

Reading Event Processing

When a reading event is generated, the corresponding page and paragraph of the content are detected, control is handed over to *Interaction Controller* (IC) module, which initiates reading event processing step. In this phase, *Annotation Manager Interface* of IC (Figure 3.4) uses the *Annotation Mapping Service* component of *Annotation Manager* module to retrieve multimedia and haptic augmentation information from the annotation XML file. This component uses the already available page ID and the matching paragraph ID information. *Annotation Mapping Service* component returns the corresponding multimedia-haptic annotation details to the *Interaction Controller* which continues to coordinate the rest of the processing.

Quiz Event Processing

Our proposed HE-Book system can be augmented with various multimedia quizzes as well. If a paragraph is previously tagged with different quiz options, then the *IC* module will be already containing corresponding quiz annotation information, and the previous reading event processing step should have been completed already. Additionally, in this stage, we need to display the packaged quiz data received from the *Multimedia-Haptic-Rendering Manager* module through the *Quiz Controller* component of the IC (Figure 3.6). As our, HE-Book system interaction is limited to paragraph level (not sentence level) hence there can be a series of quiz question-answering session and after that there will be a final score for that quiz session.

3.1.2 Annotation Manager

Annotation Manager module has three components (Figure 3.2) among which *Annotation Mapping Service* is used to retrieve augmentation information from the XML based annotation file. From figure 3.4 we see that *Annotation Manager Interface* sends {*Page*, *Para*} pair to the *Annotation Mapping Service*, and it uses that information to parse the HE-Book annotation file. By parsing the XML file, corresponding XML block is returned to the IC module which consists of Multimedia-Haptic and quiz augmentation information. Other two components *Image Search Service* and *Emotion Search Service* of this module are described in the HE-Book annotation scheme Section 3.2. We envision introducing intelligent quiz recommendation service, based on current performance of the reader in our future works.

3.1.3 Multimedia-Haptic Rendering Manager

Following section describes the details about hapto-audio-visual rendering. It also includes quiz rendering approach and how HE-Book can be integrated into a home multimedia entertainment system.

Audio Video Controller

Here, we describe the text-to-speech interaction mechanism, image-video rendering approach. *Mutimedia Haptic Rendering Interface* and *Quiz Controller* of IC sends the XML block to respective *Multimedia Data Packaging* and *Quiz Data Packaging* components (Figure 3.6) and corresponding data packaging is processed in those components.

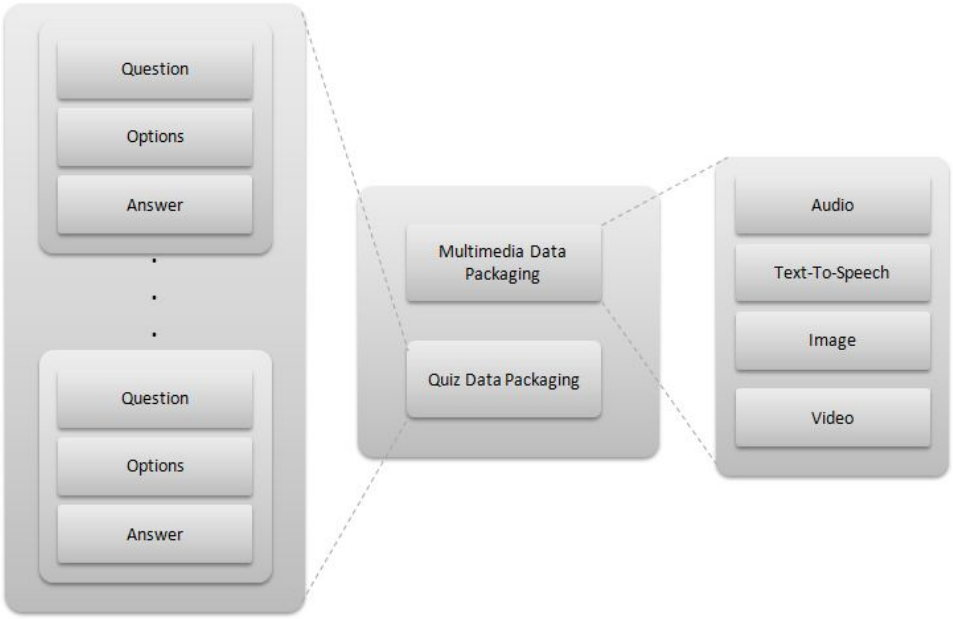

Figure 3.5: HE-Book Multimedia Packaging Sub-system

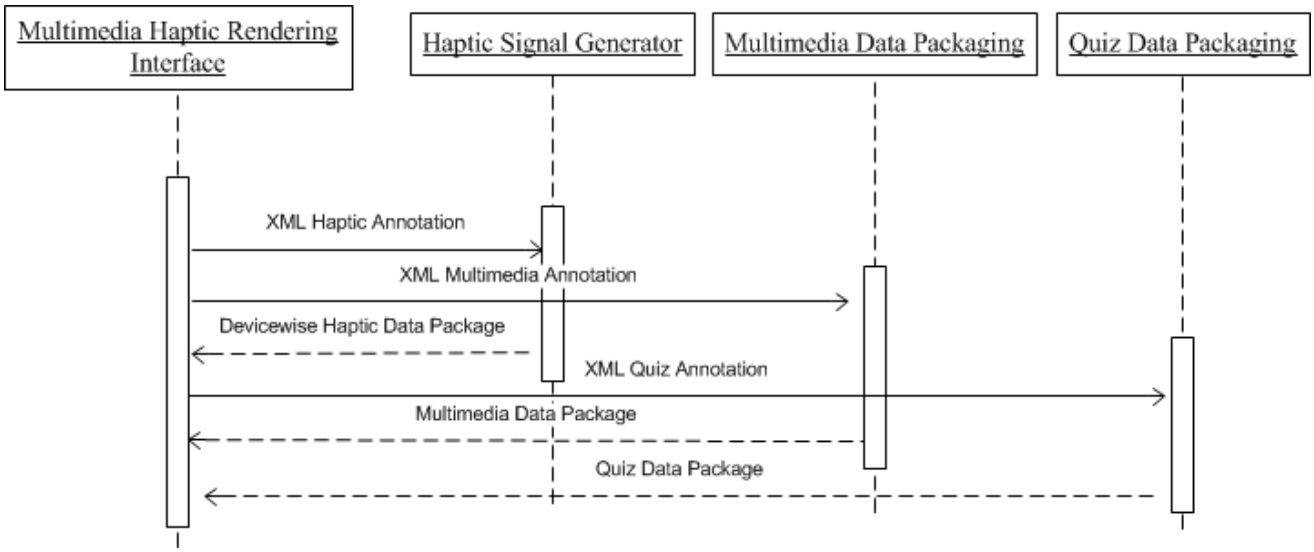

Figure 3.6: Multimedia-haptic data packaging sequence

- *Text to Speech Interaction:* In order to deliver audio feedback of the reading material to foster effective edutainment, we incorporated Text to Speech (TTS) interaction scheme in the system. Using this interaction option a reader can choose male or female voice and issue commands to listen to some selected or to all the

remaining paragraphs of the book automatically. The real time TTS processing engine produces the audio stream from the e-book content so that the learner can choose to hear the reading material. While hearing the audio content, the reader also receives the haptic rendering, and augmented images, which are delivered on the readers selected screen to 'see' the learning material by using the annotation blocks in real time. It is also possible to use the proposed HE-Book for a home multimedia entertainment system integration.

- *Image/Video Rendering Approach:* Based on the xml data received by the multimedia packaging subsystem from the IC module, the sub-system builds the media following the details given in the XML image or video block (Figure 3.5 and 3.6). For example, image/video source can be local space of the HE-Book client device, or can be coming from network place or directly from online links. Hence, image/video rendering depends on the local resources available to the HE-Book client. If the client is a simple mobile device with low hardware specifications, then this phase has limitation of usage. Therefore we assume that, the HE-Book clients will be deployed to such environment where hardware specification is not limited. For example, if we deploy the HE-Book system in a tablet device then the native imager/video viewer library of the HE-Book device incorporates augmented visual feedback to be displayed on the user device screen so that a learner can 'see' images/videos that are related to the learning content. In another scenario, if we are using a low specification mobile device as HE-Book client, then there can be an intermediate computing device which will be receiving media data following Figure 3.7 and then will playback the specific media to specific device. The later approach can be selectively used for the former approach as well. Seeing an image of a snow storm while reading about it can help the user improve his/her reading and understanding abilities in order to help practice effective learning behaviour [16]. Additionally, the augmented visual display of the learning material can make the intuitive learning process more interactive and entertaining [8].

- *Quiz Rendering Approach:* Quiz rendering process of HE-Book system follows same methodology as image/video rendering but quiz interaction is handled by the *Quiz Controller (QC)* of IC. As annotation of HE-Book is limited to paragraph level, hence each paragraph can have a set of questions and following answer (Figure 3.5). Quiz can be of many types such as Multiple Choice Questions (MCQ), fill in the blanks, crossword, anagrams, word searches etc. Here we describe MCQ only.

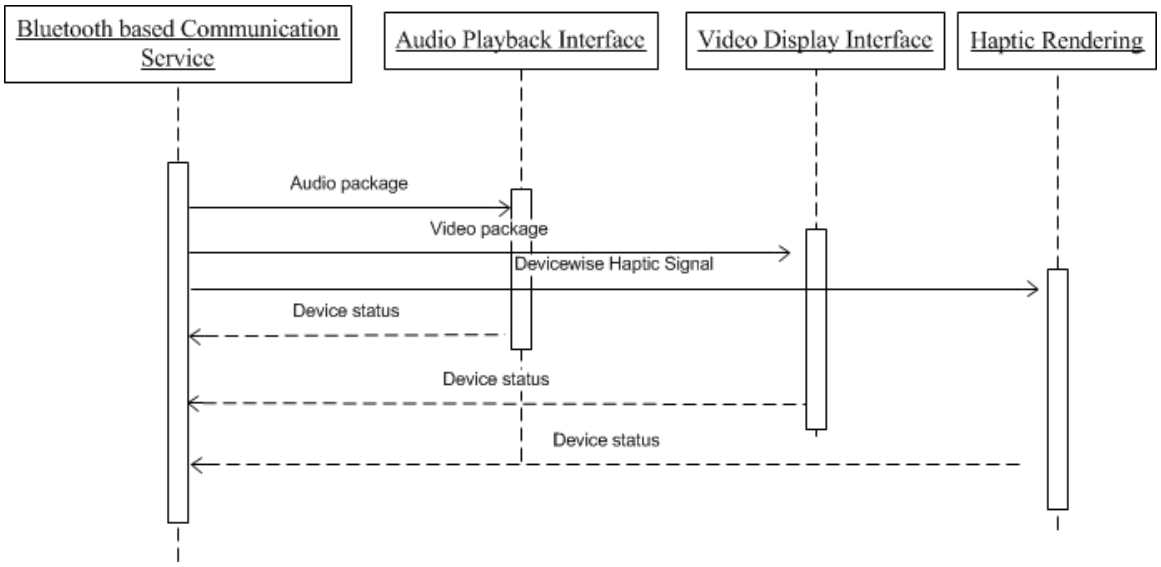

Figure 3.7: HE-Book media playback sequence in respective devices

Each quiz question is divided in the question body, multiple choice options and the respective answer of the quiz (Figure 3.8). The multiple choice option can be designed using text, image, video etc. When reading event triggers a quiz event, then the rendering is processed. Based on the xml data received from annotation manager module, *QC* displays the quiz package in target video display device following Figure 3.7. Then, it waits for the user selection of the answers and dispatches the next quiz available for this current paragraph. This process is handled by the *QC*. In our future works, we plan to add other types of quiz interactions as well.

Haptic Rendering Approach

Based on the page and paragraph information, the Interaction Controller (*IC*) receives haptic XML description from the annotation retrieving process. Next, the IC extracts haptic description blocks from the received annotation block for a given paragraph. Then, the *Multimedia-haptic Manager Interface* component of the IC manages the further processing. In this phase, haptic description block is send to the *Multimedia-Haptic Rendering Manager* module, where the *Haptic Signal Generator* component determines device specific haptic signal (Figure 3.6). As the target haptic device can be heterogeneous, we defined different types of signal for different devices. In our system architecture, we have considered Bluetooth as a method of communication between our system and the corresponding haptic device. Such communication can also be extended to other possible

Figure 3.8: HE-Book Multiple Choice type Quiz Session

personal network communication methods based on the device support. In our prototype system, we have used a Bluetooth enabled haptic jacket, a D-Box haptic sofa , and a Bluetooth enabled haptic arm-band .

- *Haptic Signal Generator:* From our previous work [2] we have selected three basic emotions {*Love, Joy, Fear*} based on the haptic devices available for our prototype system. It is also possible to extend our system to other emotion types if related devices are available. The defined emotions are tagged to the e-book content semi-automatically as will be described later. It totally depends upon the haptic devices what can be imulated through them. We have annotated real life scenarios such as *riding bike, driving car, travelling by airplane, travelling by boat, sea storm tide effects, holding hands, tickle, hug, poke, touch.*

  - *Touch:* In order to mimic touch haptic interaction we have incorporated funnelling illusion based actuator pattern generation. The patterned vibrations of the set of actuators on the human skin produce the touch sensations.
  - *Poke:* Haptic poke incorporates the similar patterned vibrations as the touch. However, the funnelling parameters for the motors are reduced and the vibration strength is higher in this type of haptic simulations.
  - *Tickle:* Through the haptic jacket interface, we empirically have positioned

the actuator motors around the belly, armpit and neck area and leveraged a set of patterned touch sensations in order to generate tickle haptic feedback.

- *Holding Hands:* By using the haptic armband we have achieved the holding hand haptic feedback. In this type of haptic rendering the haptic touch is focused on the armband area.
- *Sea Storm, Travelling in a Boat/Airplane/Car, Bike Riding:* The haptic sofa is responsible for the generation of such feedbacks. The sea storm haptic feedback is a special type of boat riding haptic feedback. In order to generate boat riding haptic sensation the D-Box interface generates wave of signals to modulate the actuators accordingly. In car travelling haptic feedback we modified the boat travelling feedback and enhanced it with frequent vibrations. In order to simulate bike-riding experience, the haptic sofa reduces the actuator vibrations, however, increases the force of the actuators in the process.
- *Fear, Joy, Love:* We again capitalize the haptic touch and haptic poke feedbacks in order to generate these emotional feedbacks. In case of fear haptic rendering, we enable the periodic patterned actuator touch sensation along the backbone. The joy effect is produced by incorporating a series of mild-poke and mild touch sensation around the stomach area which simulates butterfly effect. Whereas the love effect is produced by enhancing the haptic touch around the left-chest area. The actuator motors were placed around these said areas with varying intensity beforehand.

Home Multimedia System Integration

HE-Book system can adapt to the ubiquitous home multimedia entertainment scenario. For example, one of the HE-Book user will read the annotated content using his/her mobile device (e.g. iPad, Kindle, Google android phone) and others will enjoy the content related information by utilizing the smart devices surrounding them such as Haptic sofa, haptic jacket, haptic arm-band, stereo sound system, HDMI based TV etc. In such a scenario, the reader's partner, kids are friends are passive users. Both mom and the kid are wearing haptic jackets, sitting on a haptic sofa and in front of them, there is a HDMI enabled TV. To complete the scenario, there is also a stereo system, which can capture audio input both from the TV or a computing device wirelessly. While reading the book, an active user touches annotated part of the content and consecutively a list of images are queued to be displayed in the front TV as well as related audio can be played in

the stereo system. At the same time, wearable and surrounding haptic devices of every passive and active user will start to play the command that is sent to it. In our prototype system, haptic signalling and streaming is handled by an intermediate computer, which is connected to the mobile device using a personal communication network (e.g. Bluetooth). This scenario can be controlled by a standalone mobile device too. TTS support of the system makes it possible for the reader to relax and only listen to the machine automated reading in addition to the augmented multimedia-haptic rendering.

3.1.4 Interaction Controller

The *Interaction Controller (IC)* plays the central role to organize and synchronize the necessary work flow in the haptic e-book reading system. The IC polls to acquire user touch inputs after a certain interval. As soon as a touch interaction is performed, the IC handles the screen coordinate based paragraph identification of the e-book document. The Annotation Manager module takes the page ID and the paragraph ID to retrieve the haptic and multimedia data associated with the paragraph and returns the data to the IC. Further the IC transmits the obtained hapto-audio-visual information to the Multimedia-Haptic Rendering Manager module. In this module, haptic, multimedia and quiz information are prepared for further delivery to specific device. Device specific haptic commands are transferred to the *Haptic Signal Generator* component where haptic signal is generated for the corresponding haptic device. On the other hand, multimedia information is packaged and returned to the IC module. IC module transmits all types of multimedia-haptic data to the corresponding rendering device using personal wireless network available.

Medium of Communication

In our HE-Book system architecture, we have adopted Bluetooth as our medium of communication from the Interaction Controller module to the haptic and audio-visual rendering devices. It is also possible to use other methods of wireless communication instead of Bluetooth based on the functionality of the media devices. IC polls the surrounding active devices following the flowchart of Figure 3.9. If any media device is discovered, then it checks whether it is already a paired Bluetooth device or not. As Bluetooth pairing takes time, media ordering with minimal jitter is an important issue, so we only consider pre-paired Bluetooth devices. IC also checks the status of a device whether it is a standalone or intermediate device. Standalone device is itself capable

of continuing media streaming but for a setup such as home multimedia entertainment where number of devices are involved, an intermediate computer device play the role as mediator controller. Intermediate computer controller is introduced to overcome the computing limitations of mobile devices. In addition to coordinating the aforementioned task, IC ensures that the system modules are not blocking or throttling the operation of the other modules. The Interaction Controller employs carrier-sensing algorithm to determine the active/idle states of the other modules.

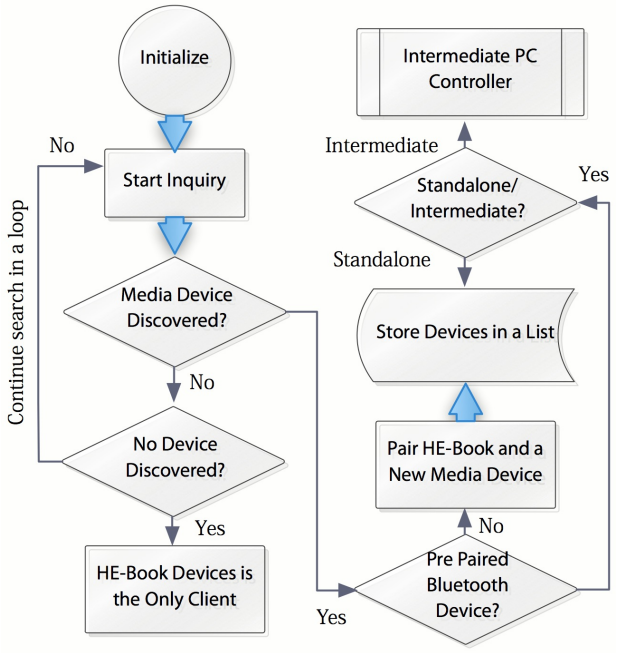

Figure 3.9: Multimedia and haptic play device discovery and listing method

Multimedia Playback Ordering

Incorporation of various media with the e-book reading scenario introduces media playback ordering issue. We have used the following Algorithm:1 to handle this issue. We have selected {*start reading, text-to-speech, image rendering, haptic rendering*} media ordering after several subjective tests using the prototype systems. Details about the result can be found in the section 5.1.2.

The media ordering algorithm runs as long as the reading event is generated. When a new a paragraph is selected by one of the available interactions, we first sense the existence of already running media stream. If already there exist running media from

Algorithm 1 Multimedia Playback Order Management algorithm for HE-Book system

Require: Client devices list from interaction controller Fig. 3.9

Ensure: Client devices are in receiving state

```
1: while Reading event continues do
2:     if New paragraph is selected then
3:         if Previous media stream is alive then
4:             Dispose haptic rendering
5:             Dispose image rendering
6:             Dispose text-to-speech service
7:         end if
8:         Start Text-to-speech if corresponding stream and device is available
9:         Start Image rendering if corresponding stream and device is available
10:        Start Haptic rendering if corresponding stream and device is available
11:    else {Media streaming is running}
12:        Continue Text-to-speech if stream and device is available
13:        Continue Image rendering if stream and device is available
14:        Continue Haptic rendering if stream and device is available
15:    end if
16: end while
```

an old paragraph, then the running media streams are disposed in the following order haptic first, image next and TTS at last. If there is no such running media, then we start to play the HE-Book media stream in the following order; TTS first, image next and haptic at last. Alternatively, if there is no new paragraph selection event, the current media streaming continues in the order it was initiated. From the subjective tests (as discussed in the evaluation chapter) we found that, if above-described media streaming order is followed, then users report minimum errors otherwise cognitive load on the user increases and surpasses his/her comfort level.

Quiz Controller

Quiz controller (QC) is a component of the IC module. It is responsible for managing all the quiz related interaction between the user and the HE-Book quiz processing system. It also maintains the paragraph wise and final scoring of a quiz session with an individual user. Basically, quiz information is stored in the XML annotation file. After the quiz information is retrieved from the annotation retriever phase, QC receives the packaged

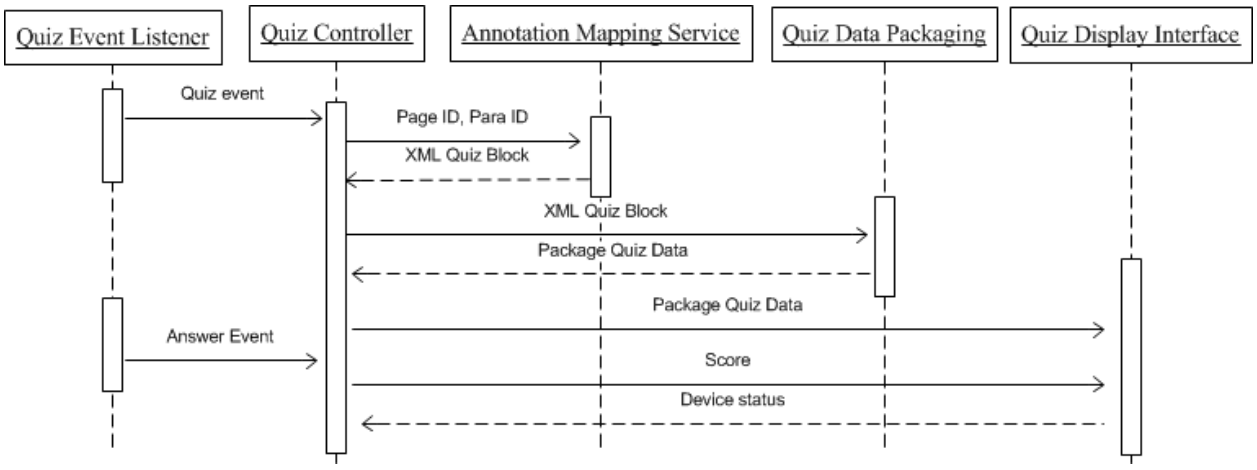

Figure 3.10: HE-Book quiz handling sequence

multimedia enabled quiz from the Quiz Data Packaging component. Next, QC displays the packaged quiz information to the HE-Book user display device (Figure 3.10). After the quiz is displayed in the user screen, user participates in the answering session and QC stores the score of the user and give feedback about the answer on the user screen (Figure3.8). As each paragraph can contain multiple numbers of quizzes, hence it is the responsibility of the QC to maintain the log of the quiz session. When, the complete book is read, QC is capable of displaying both paragraph wise and final score of the user. This system can help in tracking the learning progress of the reader.

3.2 HE-Book Annotation Scheme

In order to incorporate diverse media with current e-book architecture, we have used a separate XML based annotation file. This annotation file contains the description and details about various media (audio,video,image), haptic and quizzes. This annotation file can be extended to support other media such as the "smell" as well. We also have designed an annotation editor, in order to help in the process of annotation. Following sub sections have the detail description of the annotation language and annotation authoring procedure.

3.2.1 Annotation Language

The advantage of separate XML based annotation file (Fig. 3.11) is that, a HE-Book user could decide to either use the multimedia-haptic extension or ignore it. A possible business model for the separate XML file can be an online HE-Book XML store, from where a user can download the annotation file for a specific e-book and if necessary the corresponding media data. XML based annotation increases adaptibility to existing popular eBook file format such as EPUB. For the annotated paragraph of an e-book, we have created an XML element *Page* which contain attribute *ID* that denotes the unique number of the page. Every *Page* element is divided in several paragraphs *Para*, where *ID* denotes each paragraph no. Granularity level of our annotation is at *Para* i.e. vibrotactile feedback and multimedia playback would start and end in a paragraph leve instead of sentence level, which we termed as scene or paragraph based annotation. Alternatively we can say that the line level annotation and corresponding reading event detection is a future work that we plan to do next.

```
− <Document>
    − <Page id="1">
      − <Para id="1">
          <Haptic device="Sofa" type="Fun" duration="10" pattern="Periodic"
              Repeat="1" delay="1"/>
          <Haptic device="Jacket" type="Touch" duration="10" pattern="Periodic"
              Repeat="1" delay="1"/>
          <Home device="Speaker" type="Stereo" duration="10" Repeat="1"
              url="user/music/anthem.mid"/>
        − <Image src="Flickr" url="http://www.flickr.com/photos/1813617771/">
          <target type="TV"/> <target type="PDA"/>
          </Image>
        </Para>
      </Page>
− <Page id="2">
    + <Para id="1">
    + <Para id="3">
  </Page>
  </Document>
```

Figure 3.11: XML based annotation file for the e-book document.

Under the *Para* element, the main elements are *Haptic*, *Image* and *Video*. We support heterogeneous haptic devices as well as different type of behaviour in those devices such as Fun, Touch, Love etc. Duration and pattern of the haptic behaviour can also be customized through the XML file. *Home* element is used to incorporate home multimedia devices to the system such as sound system. *Image* and *Video* type is used to display image, video to different target devices such as Television, mobile or computer screen. It is also possible to extend the different media support to be local i.e. fetched from the

local hard drive as well as online type of *audio, video*. *Image* has attributes *Src*, *URL*. *Haptic* is described using *Device* which can be Sofa, Jacket, Arm band etc; *Type* such as Fun, Touch, Love etc; different *Pattern*, *Repeat*ation and *Delay*. Influenced by our past work [53] we annotated the e-book content with *Touch, Fun, Hug, Tickle, Kiss, Poke* type haptic feedbacks into our HE-Book system.

```
-<Document>
  -<Page id="1">
    -<Para id="1">
      +<Haptic>
      +<Image>
      +<Video>
      -<QuizList>
        -<Quiz id="1">
          <Question>What was the color of Cinderella's hill?</Question>
          <Options>
            <Option id="1" description="Red" Image="user/image/red.jpg" />
            <Option id="2" description="Blue" Image="user/image/blue.jpg" />
            <Option id="3" description="Black" Image="user/image/black.jpg" />
            <Option id="4" description="Pink" Image="user/image/pink.jpg" />
          </Options>
          <Answer>4</Answer>
        </Quiz>
        +<Quiz id="2">
      </QuizList>
    </Para>
  </Page>
</Document>
```

Figure 3.12: Quiz annotation block in HE-Book XML annotation file

For the quiz annotation part, HE-Book supports diverse multimedia quiz annotation. From Figure 3.12, we find that, every paragraph has a list of quizzes under the XML element *QuizList*. For every *Quiz*, there is an identifier *ID*, one element *Question*, set of options under *Options* list. For a quiz type of question, we can have multiple set of *Option*s. Every *Option* is identified by an *ID*. Every option can have *Description* and link url to *Image* or any other media. Next to the option set, we have the *Answer* identified by the *Option* id. It is also possible to support multiple answers by introducing a list of answers.

3.2.2 Annotation Tool

We can create annotation file following the appropriate annotation language manually or we can use any custom annotation tool for the annotation block generation purpose. In our prototype system, we have considered both manual and semi-automatic annotation procedure. Using our developed annotation authoring tool (Figure 3.13), an author can select a paragraph of a selected page of an e-book and tag it by using various haptic and multimedia properties. In case of semi-automatic annotation, the system will suggest various offline/online media resources (image, audio, video), possible emotion of the text, etc. This annotation procedure is introduced to give convenience to the annotation author.

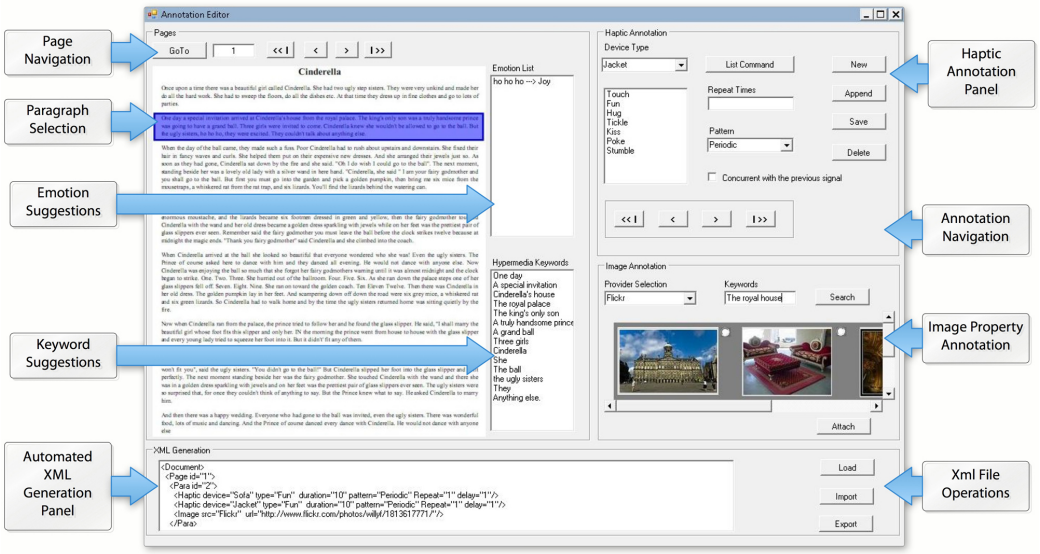

Figure 3.13: Annotation editor for the e-book document (i.e. PDF)

In the annotation authoring tool, we use Stanford POS tagger [30] to find out candidate Noun Phrase (NP) from a sentence of a selected paragraph. In order to automate the annotation process, we suggest to the annotation author a list of noun phrases, which can be good candidate for multimedia annotation. Based on the paragraph wise NP list, an annotator selects a NP that is suitable to describe a scene properly. Based on the NP selection, our image-video search service, suggests a list of online links (if available local media address) to images and videos. For the *image* annotation steps, after selecting NP from the list, the annotator is provided with a list of possible images from the available image providers based on web search and displayed them in an image container. Later,

Basic Emotion	Secondary Emotion
Love	Affection, Lust, Longing
Joy	Cheerfulness, Zest, Contentment, Pride, Optimism, Enthrallment, Relief
Surprise	Surprise
Anger	Irritation, Exasperation, Rage, Disgust, Envy, Torment
Sadness	Suffering, Sadness, Disappointment, Shame, Neglect, Sympathy
Fear	Horror, Nervousness

Table 3.1: Typed Emotion Lists for User's Personal Selection

an author selects a relevant image from the list and thus image block is attached to the annotation file. Annotator also can edit the description of the annotation block in order to better suit the purpose.

In case of haptic annotation, we have used the text based emotion detection model proposed by [3] to find out emotions of the text. In this process, the emotion detection model can suggest the author with love, joy, anger, surprise, sadness, fear after analyzing the textual paragraph. Though some emotion suggestions are ambiguous but it helps the annotator to quickly decide about possible mapping with haptic patterns. There were three supported devices: haptic jacket, haptic arm-band, haptic sofa and for each device, there is a list of supported vibration patterns. Those patterns were configured by using specific key control parameters for each haptic device. The XML file (Fig. 3.11) was further stored to be processed by the HE-Book reader.

3.2.3 Emotion Recommendation Model

In our emotion detection model, we have extended an existing model presented in Emo-Heart [42] to support detailed secondary-level emotions [48] in Table 3.1. As depicted in Fig. 2, the overall components of our proposed affective analysis system are Text Content Module, Emoticon Processor, Abbreviation Processor, Non-grammar highly intensive word processor, *EmoHeart* based Syntax, Word, Phrase, Sentence Analysis, Interjection Processor and emotion tagged output. All these modules are centrally controlled by the Step Forwarding Controller. This controller module decides which step to execute at which instance of time. Now we describe briefely each of the modules.

Text Content Module

At this phase, the text content is analyzed using Stanford POS tagger [30]. We use our predefined tags: {*emo*}, {*abb*},{*ng*}, {*g*}, {*exc*}, {*punc*}, {*int*} for this phase of tagging. Here *emo, abb, ng, g, exc, punc, int* stands respectively for emoticon, abbreviation, non-grammar, grammar, exclamation, punctuation and interjection respectively. Example of non-grammars would be loveeee, uuuu, cooool. The processed tagged input is transferred to Step Forwarding Controller.

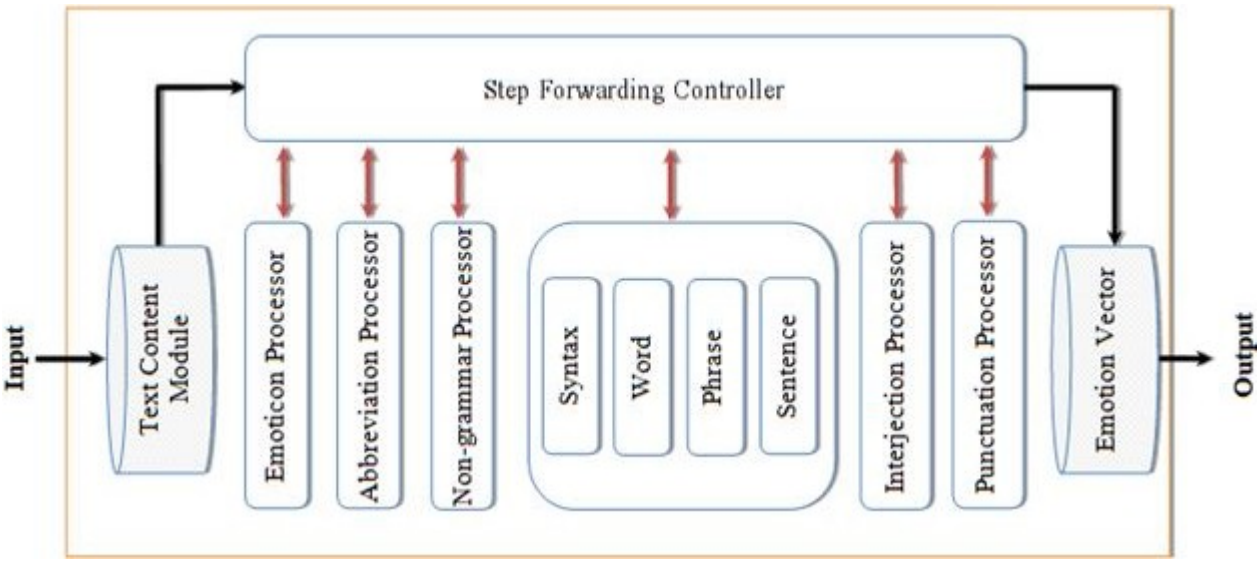

Figure 3.14: System flow of secondary level emotion detection system

Step Forwarding Controller

This module acts as the central controller as well as the transit to every intermediate input and output of the analysis model. Based on the tagged content of input (text), this module calls other low level modules to process the corresponding tagged part. At this stage, parallel execution is performed on the different tagged part of the text to improve the performance of the system. For example, emoticon processor, abbreviation processor, non-grammar processor and interjection processor can work simultaneously to translate the input text and symbols to proper transcriptions. The Punctuation processor is applied at the end of the sentence level processing. Finally, a detailed emotion description of the inputed text is obtained.

Emoticon	**Meaning**	**Emotion Label**
;-)	So happy, I'm crying	Zest
:-o	Wow!	Cheerfulness
:(	Sad	Sadness
:-ll	Angry	Rage
8-O	Omigod!!	Surprise
:-X	A big wet kiss!	Lust

Table 3.2: Partial List of Emoticons and Corresponding secondary emotions

Abbreviation	**Description**
AMBW	All my best wishes
BS	Big Smile
CUL	See You Later
DYLM	do you like me
GLLA	Great lovers love alike
HAK	Hugs and kisses

Table 3.3: Partial list of abbreviation symbol

Interjection	**Interjection**
Alas	Gee
Bingo	Hurray
Bravo	Oops
Cheers	Ssh
Dang	Wow
Eureka	Yuck

Table 3.4: Partial list of Interjection

Emoticon Processor

In this phase, the emoticon tagged contents are translated to the corresponding text emotion based on a list of secondary emotions [72]. We show a partial list of emoticons compiled on this research in Table 3.2.

Abbreviation Processor

Abbreviation processing module uses a rich abbreviation database, consisting 900 mostly used symbols compiled from [36],[1],[49],[63],[67][5]. In our algorithm, we translate one symbol found in a candidate sentence to its corresponding transcription following the compiled database. We have prepared this database after detail study of SMS, online conversation, blog and video commenting. In Table 3.3, we show a partial list of our abbreviation symbol table.

Non-grammar Processor

This module is designed to calculate a possible word from a non-grammar word. We have selected some mostly used words analyzing SMS archive [5] and public chat room texts (Yahoo Messenger, Google talk) like bye, love, u, so, cool, ah, wow etc. for this phase of calculation. When we get one of these non-grammar words hiii, byeee, loveeee, uuuu, sooo, cooooool, ahhh, woowwwww etc., we calculate its Levenshtein string distance [34] with our non-grammar basic word list. We select the basic word which requires minimum delete/insert distance from the candidate word. Also, we assume greater emotional intensity of the basic word based on the delete distance. For example, loveeeee! uuuuu has strong emotional intesity than lovee! uu.

Syntax, Word, Phrase and Sentence Analysis

We follow the steps presented in EmoHeart [42] for syntax, word, phrase and sentence analysis, but we increased the number of elements in the vector of emotional state intensities from 9 to 25. Our intention is to find more detailed emotion labels from the input text. In our model, the vector of emotional state intensity is e = *[Affection, Lust, Longing, Cheerfulness, Zest, Contentment, Pride, Optimism, Enthrallment, Relief, Surprise, Irritation, Exasperation, Rage, Disgust, Envy, Torment, Suffering, Sadness, Disappointment, Shame, Neglect, Sympathy, Horror, Nervousness]* which consists of secondary emotions. A group of this secondary emotions falls in primary emotion category

(Table 3.1).

We manually converted a small part of WordNet-Affect (539 nouns, 517 adjectives, 238 verbs and 15 adverbs) [61] to emotion intensified vector data. We have created a website which is usable in research network. Using this website (Figure 4.7), one annotator annotates the emotion intensity vector to our candidate nouns, adjectives, verbs and adverbs from the WordNet-Affect list. We have invited 3 annotators for the annotation purpose and every word was assigned to 2 annotators. Each annotator inputs emotional intensity vector of a word. When, 2 annotators suggest respective intensity vector for a single word then a conflict occurs which is then forwarded to a third annotator for conflict resolution purpose(Figure 4.8). After resolving the conflict we find emotional intensity vector of a word, such as e("compunction") is [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.1, 0.1, 0, 0, 0.1, 0.1, 0.1, 0.4, 0, 0, 0, 0.1]. Here total emotional intensity of e("compunction") is 1. Each of the 25 secondary emotions are assigned value in the range {0-1}.

In this phase, phrase and sentence level analysis follow the earlier syntax and word level analysis. After, the completion of all these analysis, we will receive a sentence level emotional vector for particular sentence of a paragraph. Next, the vector is passed to the Interjection Processing module.

Interjection Processor

This module is used at the end of sentence level analysis, because interjection changes the emotion intensity of a sentence. For example, *"I really like this dish"* is an easy calculation but when it says, *"Wow! I really like this dish."* the emotional intensity should be different from the previous one. We have selected a list of interjection words with emotional values after analyzing [55],[17],[20]. Table 3.4 presents a partial list of those interjection words. Each interjection word is annotated with emotion intensity vector neutrally i.e. irrespect of its use in a sentence, e.g. e(wow) = [0, 0, 0, 0.3, 0.2, 0, 0.1, 0, 0.1, 0.1, 0.2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]. We add the interjection vector intensity with rest of the sentence vector intensity to find the intermediate vector intesity e.g. e(wow)+e(I won the lottery). As a result, we get strong detailed emotion vector.

Finally, we have an emotion intensity vector of 25 elements for one sentence of a paragraph. From this intensity vector, we sum up the group of secondary emotions that is part of a primary emotion (Table 3.1). So, the final intensity vector is reduced to *e = [Love, Joy, Surprise, Anger, Sadness, Fear]* where *Love=Affection+Lust+Longing*,

Anger=Irritation+Exasperation+Rage+Disgust+Envy+Torment etc. From, these 6 primary emotions, we select the one which has highest vector sum as our final strong emotion for a candidate sentence.

3.3 Summary

In this chapter, we have proposed a new eBook reading platform called HE-Book, which can be annotated with diverse media (image, video, sound and haptics). The new e-book platform is adaptable in all computing platforms such as desktop, laptop, tablet, mobile. Annotation also includes multimedia quiz, which can help to track user's learning process. In order to annotate any existing ebook file format, we have introduced seperate XML based annotation file. We also present semi-automatic annotation mechanism, which helps in annotation authoring phase.

Chapter 4

Implementation Details

This chapter describes the details about various implemented prototypes of HE-Book that we have utilized for different user study. For the proof of concept, we have developed a desktop based HE-Book system, a mobile/tablet based HE-Book system, home multimedia integration, PDF annotation tool and HE-Book quiz extension.

4.1 HE-Book Prototype Implementation

We present the implementation details of the haptic e-book prototype system in Section 4.1.3. Before the implementation details we present the required system setup in Section 4.1.1. Our prototype implementation includes description about desktop based, mobile/tablet based and home multimedia system integration of the HE-Book system.

4.1.1 System Setup

HE-Book system can be used in traditional desktop based system as well as in mobile/tablet system. An user can also extend the system to work with existing home multimedia devices.

Multimedia and Home Entertainment Settings

For our home multimedia prototype, we have used Samsung LCD 46" 1080p TV for image and video display which is configured for HDMI settings. We also have a separate stereo sound system for the audio streaming. Both TV and sound system are connected with a mediator computer system. This mediator computer system is connected with

the mobile/tablet system using Bluetooth. The mediator system is later connected to the home multimedia devices using cable connection. We have used Samsung Omnia II mobile as our HE-Book reader client, which is connected using Bluetooth technology with the mediator computer 4.1.

Figure 4.1: HE-Book prototype setup

Haptics Settings

For haptic rendering we have used the following haptic devices installed or developed in our multimedia laboratory.

- *Haptic Jacket:* Our haptic jacket (Figure 4.2) consists of an array of vibrotactile actuators that are placed in location based array pattern. Location means arm, abdomen, shoulder, neck, chest etc. of a human body. Different portions of the jacket and their patterned vibration can stimulate touch in the user's skin. Vibrotactile actuators generate sound waves and create funnelling illusion when it comes into physical contacts with skin. A series of small actuator motors are placed in a 2D plane in the jacket in a certain manner. In order to produce touch feeling, the actuators are activated in a defined manner [7]. The jacket used in our prototype adhering to the aforementioned properties is discussed in detail in this work [53].

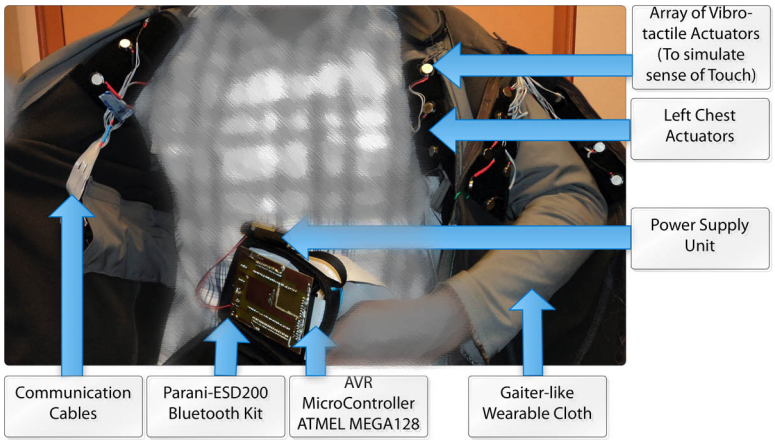

Figure 4.2: The Haptic jacket.

- *Haptic Sofa:* The D-BOX Motion Code System (Figure 4.3) is a Motion simulator system that adds the next level to the home theatre experience beyond traditional audio and visual system. D-BOX offers several models of seating with all the necessary motion equipment built-in. Add-on platforms can be attached to any chair/sofa (SRP-230 and SRP-530). The SRP-230 universal platform has four direct drive actuators that are aligned under the four corners of the sofa/chair. The actuators are capable of producing higher acceleration and lifting up to 250 pounds each (1,000 pounds in total for the four respective actuators). A control cable runs from the Motion Controller to the SRP-230 universal platform carrying the digital information that was extracted from the available media annotations. Motion controller can be of two types: PC-based model, standalone model. In our case, HE-Book mobile device system communicates using Bluetooth to the PC-based model and produces different moves using a robot event generator. We have used JAVA based robot event generator for the prototype development purposes to dynamically operate the system for lack of API to control D-BOX directly.

- *Haptic Arm Band:* Haptic arm band (Figure 4.4) works similarly to the haptic jacket system. It also has an array of vibrotactile actuators, which is a much simpler version of the haptic jacket. It has a Bluetooth communication board that takes command from Bluetooth sender device and vibrates targeted actuators in selected patterns which in turn creates funnelling illusion when comes in contact to human hands. In our case, the sender Bluetooth device is pre-installed in the mobile HE-Book device.

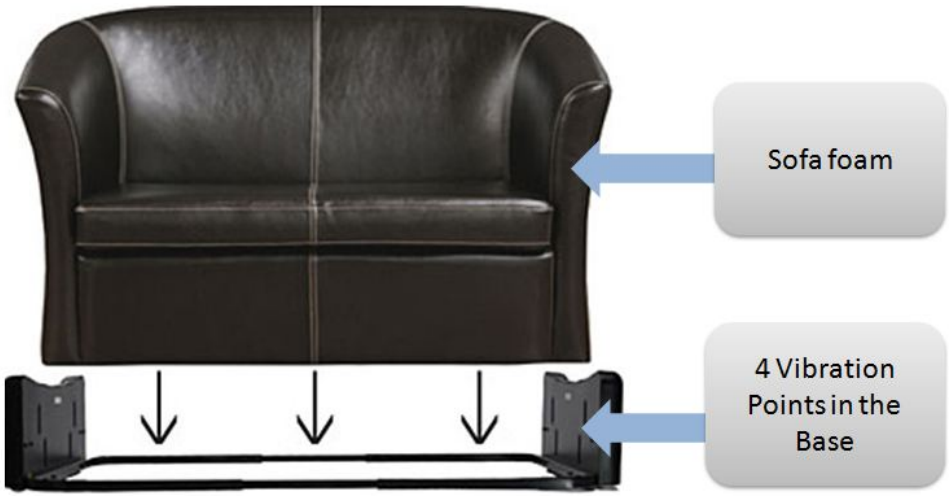

Figure 4.3: The Haptic Sofa (D-Box).

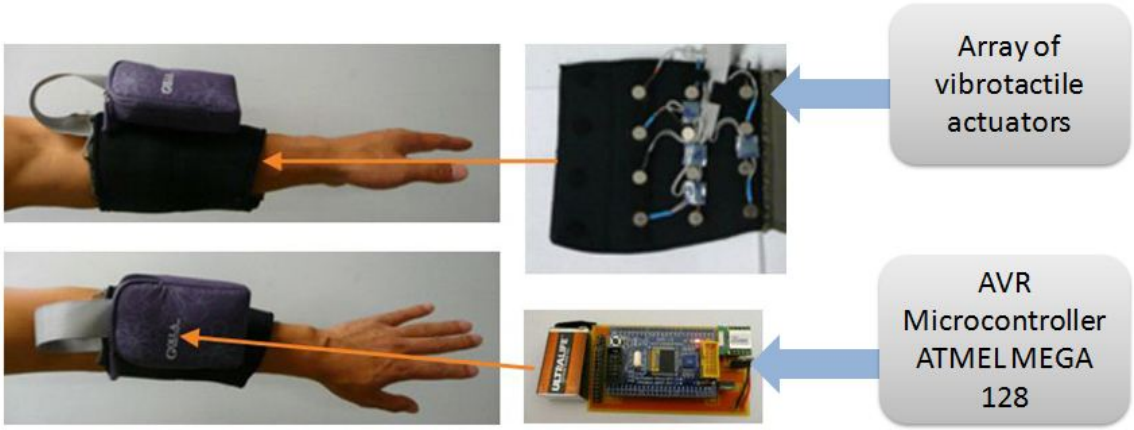

Figure 4.4: The Haptic arm band.

4.1.2 Computing Environment

HE-Book pc and mobile based prototypes are developed for JAVA environment. For the desktop version, we have used J2SE and for the mobile version we have used J2ME. The annotation authoring tool was also build in J2SE, where we have selected PDF type files as our candidate annotation target.

J2SE and PDF

Java Platform, Standard Edition or Java SE is a widely used platform for programming in the Java language. It is the Java Platform used to deploy portable applications for

general use. In practical terms, Java SE consists of a virtual machine, which must be used to run Java programs, together with a set of libraries (or "packages") needed to allow the use of file systems, networks, graphical interfaces, and so on, from within those programs [70]. Our prototypes were built using J2SE 6.0 version.

Portable Document Format (PDF) is an open standard for document exchange. This file format, created by Adobe Systems in 1993, is used for representing documents in a manner independent of application software, hardware, and operating systems.[2] Each PDF file encapsulates a complete description of a fixed-layout flat document, including the text, fonts, graphics, and other information needed to display it [71]. We have selected PDF type of documents as our candidate annotation target. Our annotation authoring tool can create XML based annotation file for PDF type of documents.

J2ME and Plain Text Reader

Java Platform, Micro Edition, or Java ME, is a Java platform designed for embedded systems (mobile devices are one kind of such systems). Target devices range from industrial controls to mobile phones (especially feature phones) and set-top boxes. Java ME was formerly known as Java 2 Platform, Micro Edition (J2ME) [69]. Our mobile prototype was MIDP 2.0 type of application.

For the mobile version of our HE-Book system, we have used plain text file reader and added our diverse media-quiz augmentation parsing facilities with it. As a result, it can act upon touch based mobile system and provide appropriate multimedia quiz on the target device.

4.1.3 System Implementation

Following section provides the details about the implemented annotation authoring tool, desktop based HE-Book system, mobile/tablet based HE-Book system, home multimedia integration method and quiz controller extension.

Desktop Based HE-Book System

In this section, we present the details of the implementation issues of our desktop based prototype system. The implemented system components are shown in Figure 3.2. For the prototype, we have developed desktop e-book readers by using Netbeans 6.5 IDE and the primary language was JAVA. In order to develop the e-book reader, we locally built the ICEPdf open source JAVA PDF viewer. We have used a Dell touch screen monitor

for our system and added the touch based interaction part to the ICEPdf source. We also have implemented XML retrieving and matching facilities to the ICEPdf for annotation searching. For serial port based Bluetooth communication we have used GPL based BlueCove[1] library suitable for J2SE based Bluetooth communication. Our prototype system communicates with a haptic jacket (Figure 4.2) when any annotated e-book part is touched by the user. A Bluetooth device was connected with the PCs USB port, which was virtually configured with the COM port so that the Bluetooth device can send signals to the haptic jacket. Our desktop based test prototype was adequately responsive in a standard Pentium dual core 32-bit machine with 2 GB systems RAM. In order to annotate an e-book, we have taken PDF as example and developed a PDF annotator modifying the ICEPdf that we have used as a viewer also. For the text-to-speech support, we have used FreeTTS[2] a speech synthesizer written entirely in the Java. By default, this application uses the *kevin16* voice that comes with the *lib/cmu_us_kal.jar* file. One can, however, change a voice by passing the name of a voice to the engine.

Mobile and Tablet Based HE-Book System

In order to develop the mobile version of the haptic e-book reader, we selected Samsung I8000 Omnia II (Figure 4.1) Windows mobile phone as the deployment device. We have developed our mobile system in J2ME using Netbeans IDE. The SDK provided the basic packages to retrieve and display images on the mobile screen. The touch APIs were available both to determine the screen coordinate and mapping of the paragraph IDs of the touched texts from an e-book XML file. We used RFCOMM based Bluetooth Serial Port Profile communication with the end haptic devices. Our end Bluetooth devices were haptic jacket and haptic armband. The command format was suitable for sending/receiving Bluetooth signals to the haptic devices through Parani ESD200 Bluetooth kit. For text-to-speech version of the mobile system we have used *java.speech* package under JSR 13 Java Speech API.

Home Multimedia Integration

For the home multimedia integration prototype, we have combined the above described two systems in one system where desktop system acts as mediator server and the mobile device as a client. When a reading event is initiated from the Samsung Omnia II mobile

[1]BlueCove, http://www.bluecove.org

[2]FreeTTS 1.2, http://freetts.sourceforge.net/docs/index.php

Figure 4.5: Hypermedia based e-book application deployed in a mobile device

device, a specific command is transmitted to the desktop server system using Bluetooth. Intermediate desktop system receives the command and distributes various media to the corresponding device (such as image-video to TV, audio to stereo, haptics to haptic devices) in the home multimedia settings. For the home multimedia version, we had the option to stream related pictures from the mobile device as well as from the desktop system. In the case of mobile based streaming, picture data is received in the desktop end and displayed on the Samsung LCD 46" 1080p TV, which is configured for HDMI display. On the other hand, the desktop based streaming loads images from the hard disk or from internet by mapping the identification provided by the mobile device and display on the TV. Mobile device retrieves the image identification from the annotation XML file that resides in the mobile device.

Quiz Controller

The quiz controller was implemented using J2ME for the user interface creation purpose. When, one user selects a paragraph augmented with quiz sets, the quiz session starts like a pop up and expects the user response through selection. Based on the appropriate selection of the user, answer is matched with the stored XML information, and users score is displayed in the mobile/tablet screen (Figure 3.8). XML matching is done on the mobile device. Mobile system tracks the final score of the user-quiz session.

Annotation Authoring Tool

For the prototype, we have developed desktop based e-book annotation tool (Figure 3.13) using Netbeans 6.5 IDE and the primary language was JAVA. For this purpose, we have

locally built the ICEPdf [3] open source JAVA PDF viewer. We have considered PDF as example e-book. We have added tagging window in the system where we have considered various tagging options. An annotation author can navigate between the PDF pages and select any paragraph to annotate.

For every paragraph, a list of hypermedia keywords motivated by Noun Phrase detection is provided to help the annotator. He/she can select the NP's and search for possible image/video/audio from online provider websites. Following these steps, an annotator can create image, video annotation XML block for the annotation file. For every paragraph, we also provide a list of emotions possible which guide the annotator to create haptic blocks. Haptic scenes such as riding a horse/boat are undetectable and needs the attention of the annotator for creating haptic block. Scenary detection and suggestion can be a good future work. Our editor supports various list of haptic devices and their customized annotation. Image and video are searched using web services in the famous image, video web stores and are shown in a panel for author's selection. For each tagging, an XML block is created, which is further editable by the human operator for various haptic attributes. Later, this XML file is stored to be further used by the ICEPdf viewer while reading annotated e-book. If any local (not internet) media is used as the annotation target, then it is also packaged with the XML file in a separate folder.

For the MCQ type quiz annotation, an author selects one paragraph and adds the quiz details in the quiz creation panel. It includes the question and number of MCQ options to be created. Every MCQ option, can be tagged with textual description and/or multimedia information (Figure 4.6).

- *Image-Video Augmentation Recommendation Service* We have used the open source *The Stanford Parser: A statistical parser* for parsing our paragraphs and detecting appropriate Noun Phrases, which are later used to search online Image and Video databases such as Flickr, Google Image Search etc. Top 5-10 results from the search results are suggested to the HE-Book annotation author.

- *Emotion Recommendation Service* We have implemented emotion recommendation service using JAVA technology. In order to annotate part of the 539 nouns, 517 adjectives, 238 verbs and 15 adverbs of WordNet-Affect [61] we have created a website which is usable in research network. Using this website (Figure 4.7), one annotator annotates the emotion intensity vector to our candidate nouns, adjectives, verbs and adverbs. We have invited 3 annotators for the annotatin purpose

[3]Open Source Java PDF, http://www.icepdf.org

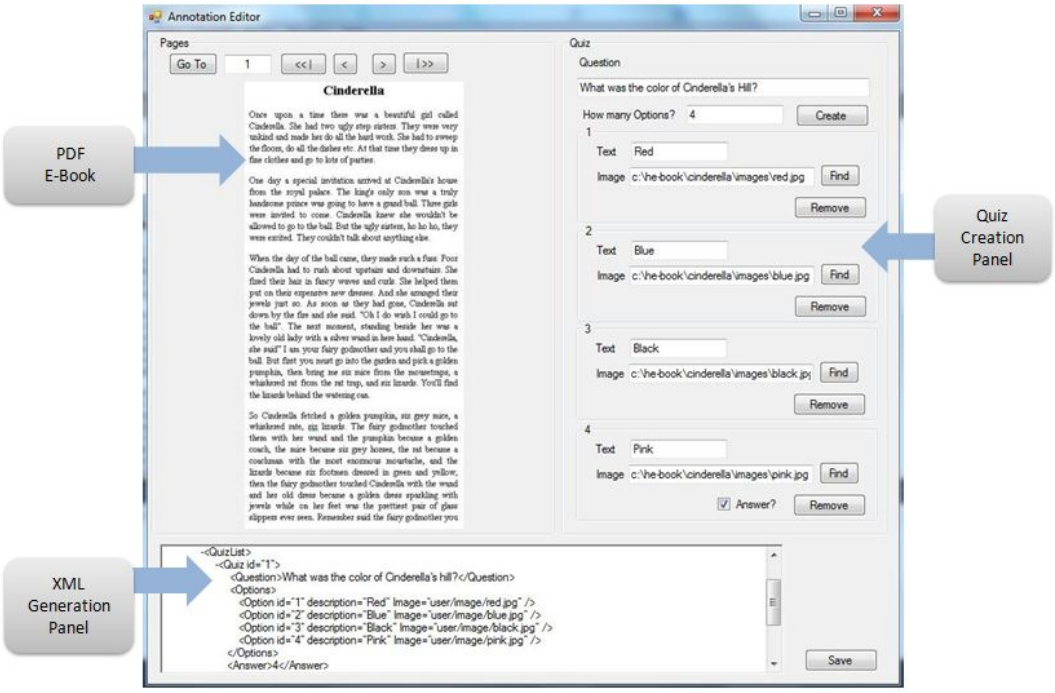

Figure 4.6: HE-Book quiz annotation editor

and every word was assigned to 2 annotators. Each annotator inputs emotional intensity vector of a word. When, 2 annotators suggest respective intensity vector for a single word then a conflict occurs which is then forwarded to a third annotator for conflict resolution purpose(Figure 4.8). After resolving the conflict we find emotional intensity vector of a word, such as e(compunction) is [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.1, 0.1, 0, 0, 0.1, 0.1, 0.1, 0.4, 0, 0, 0, 0.1].

In our system, the emotion recommendation system is hosted in a server which is accessible using network and we send the request using web service to the emotion analysis model from the HE-Book host and the server finally returns the recommended emotion list for a selected paragraph.

4.2 Summary

In this chapter, we have described the details about several prototype implementations of the proposed HE-Book system. We have implemented desktop, tablet/mobile based HE-Book system as well as desktop based HE-Book annotation tool. We also imple-

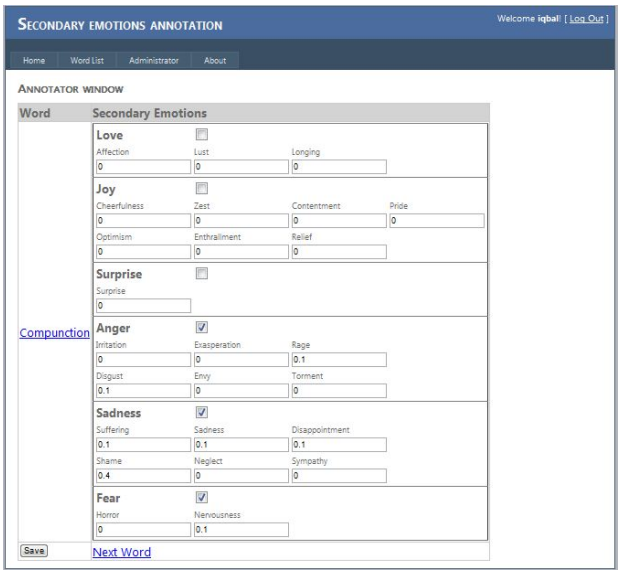

Figure 4.7: Secondary Emotion Annotation Window of the Annotator

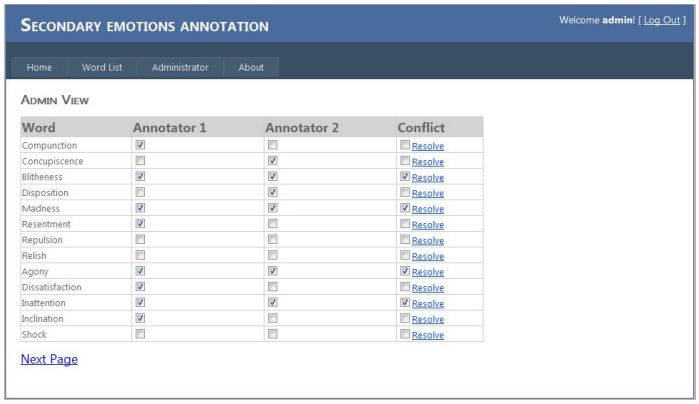

Figure 4.8: Emotion annotation conflict resolution

mented the semi-automatic annotation phase where our annotation editor can suggest related noun phrase and emotions by analyzing the text content, which is further used for multimedia annotation. We also implemented HE-Book system extensions such as quiz augmentation and ubiquitous home entertainment setup.

Chapter 5

Results and Discussions

In order to evaluate the performance and user acceptance of the he-book system, we have conducted various qualitative (in Section 5.1) and quantitative (in Section 5.2) measurements. We also conducted indirect objective based user study (in Section 5.3) to find the effect of multi-modal augmentation to e-book reading. While presenting the results, we have provided necessary discussion about the results alongside.

5.1 Quantitative Measurements

The following sections describe the quantitative measurements carried out in the laboratory in restricted environment. Observations and analysis of the results are also given with the results. These results are acquired using the system setup described in earlier chapter.

5.1.1 Haptic Rendering Delay

In order to calculate the total haptic rendering delay, we also have to consider the method of personal communication network we are using. For our prototype system, we have used Bluetooth to communicate from the mobile HE-Book device to the intermediate computer device. We have used serial port profile based RFCOMM protocol[1] for the communication. RFCOMM sends the user data (packages in frames) to the lower layers of Bluetooth stack via L2CAP (Logical Link Control & Adaptation Protocol)[40]. According to the study of [40], in ideal condition, if no Bluetooth packet is lost in the

[1] http://en.wikipedia.org/wiki/Bluetooth_protocols

communication, then delay is negligible, and this delay does not depend on the USB port or the operating system. However, if Bluetooth communications are exposed to interference of other Bluetooth, 802.15.4 (Zigbee), 802.11 (WiFi) interfaces or any other device with proprietary short-range communications protocol working in the same ISM band, a minimum Bluetooth communication delay can be assumed. Here, our HE-Book client is a mobile application, and it is connected with a mediator computer through Bluetooth. Our simulation event starts when we touch HE-Book paragraph and end when respective three haptic devices are played. We have projected the time difference between these two-time events.

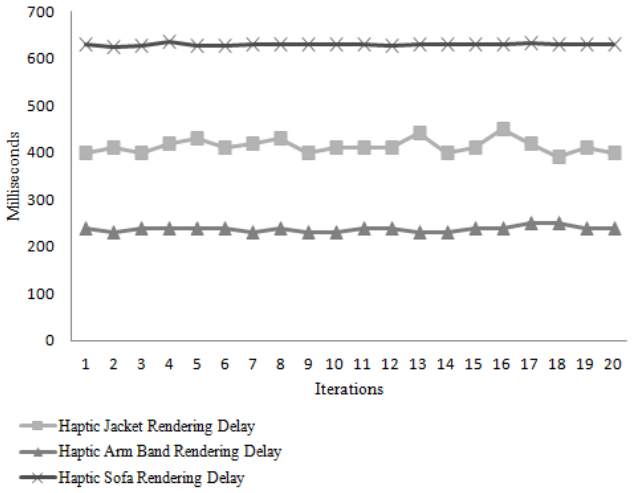

Figure 5.1: Haptic rendering delay for different devices

From the above graph (Fig. 5.1), we conclude that, haptic arm band rendering shows the least delay, followed by the haptic jacket, and finally the haptic sofa. The arm band consists of fewer numbers of actuators and requires fewer numbers of haptic Media Unit (MU) bytes hence the delay is low. Whereas for jacket, MU is relatively higher than the arm band and the number of actuators is also high. As a result, it takes more time than the arm band haptic rendering. In case of a haptic sofa, we observe a higher delay as it has longer MU bytes and heavy vibration capable actuators. But on the average, the overall delay is tolerable for this kind of haptic devices [50]. This observed delay is important for meaningful media playback and their respective ordering.

5.1.2 Multimedia Perception Error

Diverse media augmentation and reading activity introduces proper media ordering and playback challenges in order to maintain satisfactory cognitive load on the user. So, we need to find a comfortable media playback order and acceptable media skew. For this purpose, we have conducted a subjective test, where a user experience same reading event following different random media order. If a user reports a problem, while following a reading activity augmented with the random media playback order, then we record synchronization error. In Figure 5.2, we show the error rate as a function of the above defined synchronization error. We have adopted the method of [58], where error rate is defined as the ratio of the number of subjects who perceived the synchronization error to the total number of subjects. Synchronization error denotes the skew among various media. When the synchronization error is positive, a media is started after the 200 words/scene reading event started. When the synchronization error is negative, a media is started earlier than the 200 words/scene reading event started. We have compiled 200 words based scene paragraph because according to Carver et al. [14] college students generally operate their comprehension reading process at rates around 200 words per minute, and all of our test users fall into this age group. 10 college students participated in the test, where each user volunteered around 30 minutes to observer different media order. Finally, the result is plotted in the graph of Figure 5.2.

From Figure 5.2, we conclude that for a certain media playback ordering of {*start reading, text-to-speech, image rendering, haptic rendering*} with positive synchronization error $20ms$ *to* $260ms$, the error rate is less than 20%. That is people hardly reported synchronization error for such media distribution and play time difference. However, if the order of the presented media changes such as *text-to-speech, start reading, haptic rendering, image rendering* or any other, except the 1st presented order, the error rate starts to increase, that is people report more synchronization error. It is also an important observation that, it is possible to omit one media from the 1st described order but still continue with the low error rate. For example, if {*start reading, text-to-speech, haptic rendering*} or {*start reading, image rendering, haptic rendering*} is followed where image rendering and text-to-speech is omitted respectively, still synchronization error rate is low. So, in our multimedia rendering manager we have handled the synchronization issue according to the above mentioned media playback pattern in media skew range $20ms$ *to* $260ms$ using the Algorithm 1.

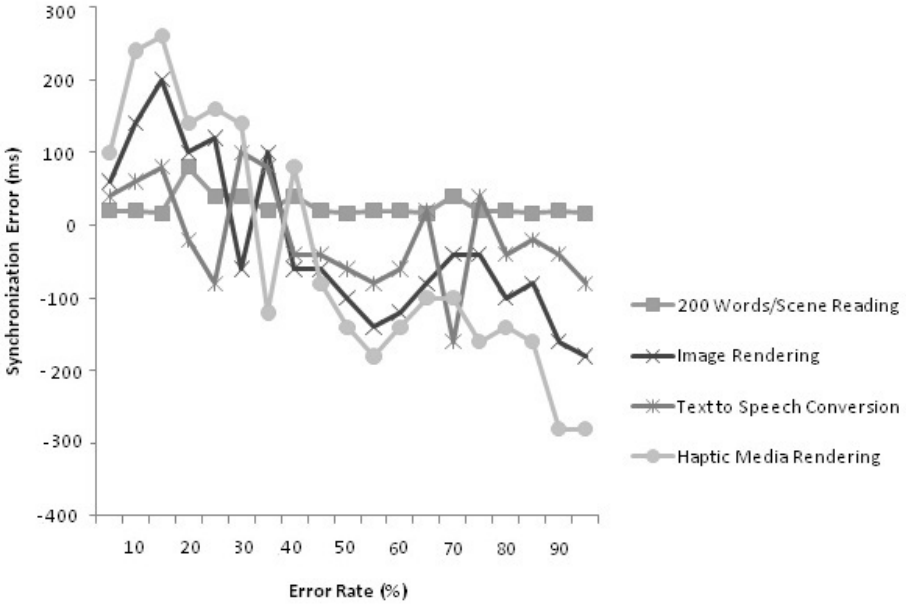

Figure 5.2: Error rates of various media versus synchronization error

5.2 Qualitative Measurements

Following sections describe the qualitative measurements carried out in the laboratory in restricted environment to determine the usefulness of the proposed system.

5.2.1 Haptic Rendering Realization

After empirically defining the haptic sensations, we have devised usability experiments to understand how the users perceive the defined haptic feedbacks. Moreover, we conducted usability tests to evaluate the user's quality of experience with our HE-Book system and to justify the suitability of the proposed approach. The usability tests took place at our laboratory in a controlled environment with 20 volunteers of different age groups, sex and haptic experience. We have selected 2 groups of users, one group includes users who experienced haptics while the other group includes those who did not. The users were requested to use the desktop-based prototype and were given PDF documents to read, which were previously annotated with haptic, audio and visual data. Their activity were monitored throughout the experiment and noted for analysis. Afterwards, based on their interaction experience, they were asked to rate various haptic feedbacks related to visual experience in a Likert Scale. The ratings of the feedbacks were in the range of

1-5 (the higher the rating, the greater is the satisfaction). In Figure 5.3, we depict the user perception result of the available haptic events based on the similar questions asked in [33].

1. I felt myself "drawn in"
2. I enjoyed myself
3. My experience was intense
4. I had a sense of being in the book scenes
5. I responded emotionally
6. I felt that all my senses were stimulated at the same time
7. The system is easy to get familiar with
8. I would consider using this system.

From the empirical study, we observed that most of the users agreed about the haptic effect combined with visual experience except for *travelling aeroplane*, and *fear*. Our observation is, proper representation of scenarios and emotions depends upon the capability of the individual haptic device and there is always scope to enhace quality of experience.

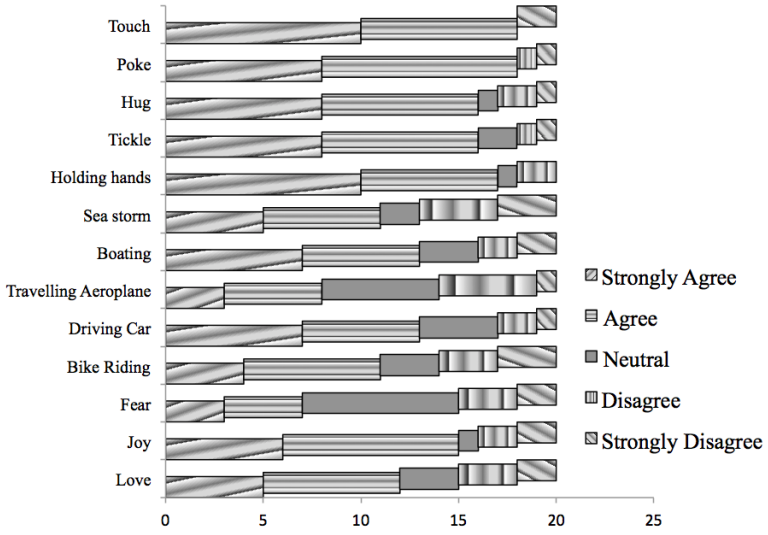

Figure 5.3: User response to various haptic feedback in Likert Scale

5.2.2 Quiz Augmentation in E-Book

We have prepared some questionnaires for users other than the previous tests to collect their opinion about augmenting quiz to the e-book platform. In this case, we have asked them few questions *"Quiz Augmentation is a good idea"*, *"Quiz will help to maintain attention in reading activities"*, *"Quiz augmented E-Book will be especially helpful for young students"*, etc. and requested them to provide their rating in scale of 5. Most of the users agreed that quiz augmentation in E-Book platform can be widely accepted by the usual users (see Figure 5.4).

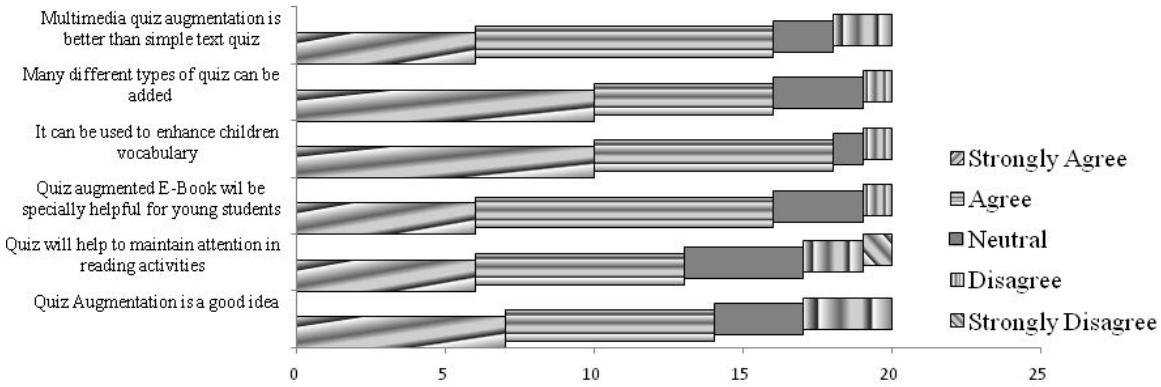

Figure 5.4: User interest in e-book based multimedia quiz augmentation system

5.3 User Study

Indirect objective based performance analysis is the most commonly used method of investigation in multimedia learning [10]. We have adopted this technique to investigate the comparative effect of each media on multimedia reading process. During the usability experiments, we monitored the activities of the users and noted those for further analysis. After completion of the experiment with a user, we have provided respective knowledge acquisition scores and requested the user to rate three assertions in Likert Scale [35]. The ratings of the assertions were in the range of 1-5 (the higher the rating, the greater is the satisfaction). The assertions were A1) The modality used in the experiment instills attention in the reading process, A2) The feedback modality improves the retention process, and A3) The modality makes the learning process more entertaining. Outcome measures of the learning task are objective because they measure performance. They are indirect because they depend on processes of information storage and retrieval that may

be affected by cognitive load. The typical design of research studies using this approach compares two or more different variants of multimedia instruction of the similar material. The amount of intrinsic loads induced by each of the variants is presumed equal because the informational contents of these materials are somewhere identical. It is, therefore, assumed that the more knowledge the learners acquire, the less extraneous load is induced by the content.

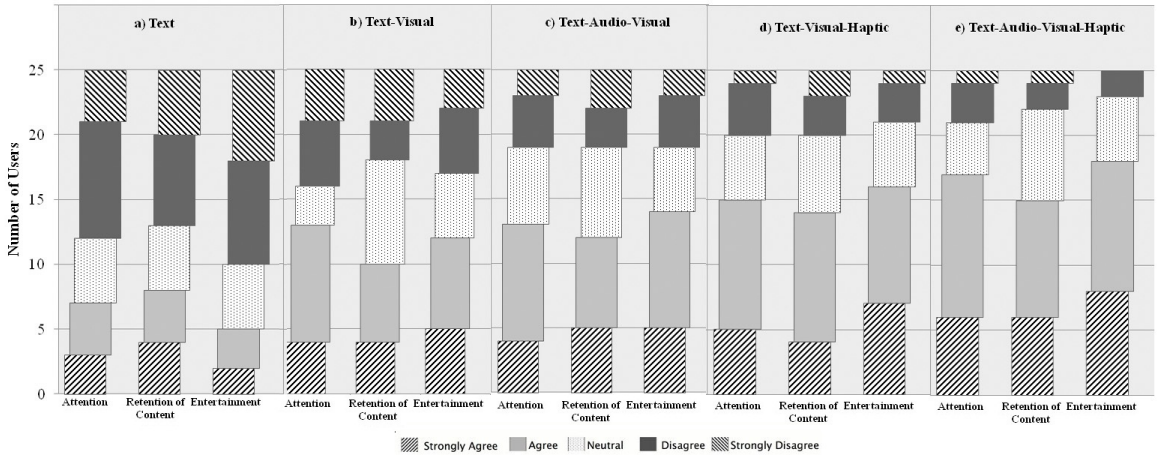

Figure 5.5: Usability study of the HE-Book system. a) Likert scale presentation of the attention, rentetion of content and entertainment metrics in a traditional text reading experience, b) The same study with text-visual rendering added in the reading process, c) Incorporating audio-visual simultaneously to determine the changes in reading process d) Visual-haptic feedbacks added in the reading process, e) Combining the hapto-audio-visual feedbacks in the paragraph reading experience to report the user scores.

To assess the influence of various levels of augmentation on the user, we selected fairy tales (*Cinderella*, *Beauty and the Beast*, *The Frog Prince*, *The Golden Goose*) as our candidate annotation material. The fairy tale is packed with adventure, fun, love, etc. in few paragraphs, it is also easier for a user to complete it in a small amount of time. For each of the fairy tales, we have compiled a new version introducing fine details inside the story. For example, we have added the color of Cinderellas dress, color of her hill, her hair color, details about the ball room, details about how prince look like, etc. These details are used to test users retention of the content.

To understand user's learning behavior, we have set up five test scenarios for each user. The tests were taken place at the university laboratory in a controlled environment where 25 (17 male, 8 female) volunteers participated. They were from 12-18 age group, and

little to no haptic experience. In each case, the participating user was given a fairy tale consisting of five paragraphs to read and at the end to answer two MCQ questions from each paragraph totals to ten. The interaction modalities in each test-bed were targeted to determine the *Attention, Retention and Entertainment* metrics of a user during his reading activity. The cumulative findings from each test case are depicted in Figure 5.5. In the first test A, users were asked to read a paragraph only consisting of textual information. In the second test B, in addition to textual information, we annotated visual data to another paragraph so that while reading certain section of the paragraph, the user can see related scenes. Next in C, another paragraph was read out loud by using the TTS and corresponding visual scenes were rendered to the screen. In the experimental setup of D, a new paragraph was annotated with haptic feedbacks and augmented image was rendered on the screen that explained the learning material. In the final setup E, one paragraph was annotated with haptic, audio and augmented image material. Paragraph wise media annotation set is not sequential rather randomly distributed. Also, each user is randomly given choice to choose among the list of fairy tales. We have designed these test scenarios to differentiate the effect of one media from other where {*text, text-visual*}, {*text-visual,text-audio-visual* }, {*text-visual, text-visual-haptic*}, {*text-audio-visual, text-visual-haptic*}, {*text-visual-haptic, text-audio-visual-haptic*} are possible sets for effect comparison. After completion of the tests with one user we handed him his knowledge score from MCQ questions.

In our experiments, we found that the well known *cone-of-learning* is a true effect in the e-book reading process. In our tests, we have considered *Strongly Agree, Agree* as positive score and *Disagree, Strongly Disagree* as negative score. Overall positive score of text-based learning was 27% in our experiments, which is improved to 47% when any visual media is augmented with the text content. We find that the effect of audio augmentation to any media mostly increased the score overall by 6%, whereas addition of haptic media increased the overall score by 13%. Possible reason for a high score while receiving haptic feedback could be due to the excitement that the users felt in receiving haptic feedbacks that relate with the learning materials. As visual feedbacks by itself improves memorization, hence the comparison with *text-visual* and *text-visual-haptic* shows that haptic also has an effect to retention. For the combination of *text-audio-visual-haptic* media with previously mentioned media order, overall score reached maximum 67% which is mostly 15% more than *text-visual* or *text-audio-visual* based media augmentation. Hence we find that, HE-Book system upholds the cone-of-learning phenomena and *text-audio-visual-haptic* media annotation ensures maximum

benefit from e-book annotation system.

5.4 Limitations of the HE-Book System

Current reading event granularity level is limited to paragraph level that is; multimedia-haptic augmentation playbacks work at the paragraph level not in fine detail sentence level. Also, our current reading event generation procedure is dependent on human finger touch interaction or automated machine reading settings. Both are not purely intuitive interaction method to existing e-book platforms. This interaction can be upgraded by more intuitive eye gaze based interaction or gesture based interaction which we consider as future works. For the semi-automatic annotation procedure, we need to conduct a large-scale performance evaluation of the affection detection model. In addition, user study needs to be performed in real-life environment other than laboratory only. This will help to find the true cognitive loads on the normal user. In our system setup, we have made calculations based on mediator computer, but we need to do standalone tablet to surrounding device communication as well.

5.5 Summary

In this chapter, we have described and discussed the evaluation results we have recorded while testing our prototype systems vy volunteers of different age, sex and multimedia-haptic experience. We have presented several quantitative and qualitative measurements on the prototype systems. Our measurements show that, HE-Book system can be a good addition to existing eBook reading experience and there is a positive tendency of accepting such a system. As our experiments are conducted at laboratory setup with a limited number of users, we target to conduct real life large-scale tests in future to find the actual effect of our proposed system on eBook reading experience.

Chapter 6

Conclusion and Future work

6.1 Conclusion

Human computer interaction research has gained a pace with the advent of haptic devices. Our research goal was to explore the possibilities and opportunities this modality could bring to our traditional eBook reading experience. In this regard, we have conducted existing literature review and found that multimedia-haptics augmented eBook platform can be a strong addition with current augmented reality based books as eBooks are replacing real books in many cases.

In this research, we have introduced an intuitive e-book reader, which is capable of providing hapto-audio-visual interaction to user's eBook reading experience. We have presented a generic architecture for diverse media augmented e-book reading system as well as described the detail component architecture, including sequence of interactions. Though we have described the system targeting text, audio, video, haptic, but it can be extended to other media such as smell. Our generic architecture is capable of adapting new extensions such as multimedia quiz and ubiquitous home entertainment setup as well. Our proposed system can meaningfully synchronize various media to corresponding devices to form a better orchestra of learning and entertainment. For this purpose, we have devised an algorithm which can order the different media playback maintaining user comfort.

For the annotation purpose, we have developed open source based PDF authoring tool. We also developed several prototypes of the proposed system extending open source PDF reader and have evaluated the performance, acceptance of the complete system using various measurement parameters. We have made five controlled usability tests to

determine reader's experience to multimodal reading activity in the domain of attention, retention and entertainment. Various analyses of the results show that, our system can help the reader to immerse into the reading scenario which in turn influences positive improvement over traditional e-book reading.

6.2 Future Work

In order to test the performance of the system and to collect various user interactions related data, we have established controlled setup in a laboratory environment. But, to make concluding remarks about the educational development of the users, it requires large scale real world tests for a significant amount of time. In future, we plan to conduct more real world experiments with our HE-Book platform. As we have adopted XML based annotation methodology and most widely used eBook format is *EPUB*, hence we want to contribute in the EPUB Media Overlay upgrading to support video, 3D, haptic etc.In future research, we can add other modalities such as smell and experiment with the user response.

As our existing system supports annotation granularity up to eBook paragraphs rather than the detail sentence level, it can be extended to eye gaze based intuitive interaction, which will enable it with detailed annotation as well as hands free communication. Our semi-automatic annotation procedure can also be upgraded to support greater level of precision. For this purpose, we can adopt the statistical analysis instead of rule based analysis. To measure the performance of the algorithm, we need to annotate more WordNet-Affect words, as well as, we have to conduct large scale data analysis.

It is also possible to extend the system to multiple peer-to-peer HE-Book communication, where a group of, HE-Book readers can participate and collaborate in eBook reading session. A potential example of such application can be, grandparents reading out eBooks from remote locations to the grand children. We also can add additional highlighting or bookmarking features to the eBook system.

We hope that, this research will attract attention of other researchers to explore the opportunities of augmenting diverse media with e-book materials as well as incorporating ubiquitous system support with them.

Bibliography

[1] Abberviations. Abberviations @ONLINE. May 2010.

[2] K. M. Alam, S. Hardy, A. Akther, and A. El Saddik. Sms text based affective haptic application. In *Proceedings of Virtual Reality International Conference (VRIC 2011), To be published*, 2011.

[3] K. M. Alam, A. S. M. M Rahman, A. El Saddik, and G. Wail. Adding emotional tag to augment context-awareness in social network services. In *Instrumentation and Measurement Technology Conference (I2MTC), 2011 IEEE, To be published*, 2011.

[4] V. D. Andries and E. R. David. Computers and publishing: Writing, editing and printing. In *Advances in Computers 10*, pages 145–174. Academic Press, 1970.

[5] SMS Archive. Sms archive @ONLINE. May 2010.

[6] R.T. Azuma et al. A survey of augmented reality. *Presence-Teleoperators and Virtual Environments*, 6(4):355–385, 1997.

[7] A. Barghout, C. Jongeun, A. El Saddik, J. Kammerl, and E. Steinbach. Spatial resolution of vibrotactile perception on the human forearm when exploiting funneling illusion. In *IEEE International Workshop on HAVE*, pages 19 –23, 2009.

[8] M. Billinghurst, H. Kato, and I. Poupyrev. The magicbook - moving seamlessly between reality and virtuality. *Computer Graphics and Applications, IEEE*, 21(3):6 – 8, 2001.

[9] S.A. Brewster, F. Chohan, and L.M. Brown. Tactile feedback for mobile interactions. In *Proceedings of ACM CHI*, pages 159–162, San Jose, CA, USA, 2007. ACM Press/Addison-Wesley Publishing Co.

[10] R. Brunken, Jan L. Plass, and D. Leutner. Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, 38(1):53–61, 2003.

[11] S. K. Card, L. Hong, J.D. Mackinlay, Ed. H. Chi, and Chi H. 3book: A scalable 3d virtual book. In *In Extended abstracts of the 2004 conference on Human factors and computing systems (CHI*, pages 1095–1098. ACM Press, 2004.

[12] S.K. Card and D.A. Henderson Jr. Catalogues: A metaphor for computer application delivery. In *INTERACT'87*, pages 969–964, 1987.

[13] S.K. Card, G.G. Robertson, and W. York. The webbook and the web forager: an information workspace for the world-wide web. In *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground*, pages 111–ff. ACM, 1996.

[14] R. P. Carver. Reading rate: Theory, research, and practical implications. *Journal of Reading*, 36(2):pp. 84–95, 1992.

[15] N. Chen, F. Guimbretiere, M. Dixon, C. Lewis, and M. Agrawala. Navigation techniques for dual-display e-book readers. In *Proceeding of the SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1779–1788, USA, 2008. ACM.

[16] E. Dale. *Audio-Visual Methods in Teaching*. Rinehart and Winstor, 1969.

[17] Your Dictioanly.com. List of interjections @ONLINE. May 2010.

[18] M. Eid, J. Cha, and A. El Saddik. Hugme: A haptic videoconferencing system for interpersonal communication. In *Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2008. VECIMS 2008. IEEE Conference on*, pages 5–9. IEEE, 2008.

[19] A. El Saddik, M. Orozco, M. Eid, and J. Cha. *Haptics Technologies: Bringing Touch to Multimedia*. Springer, 2011.

[20] EnglishClub.com. Interjections @ONLINE. May 2010.

[21] E. Gardiner and Ronald G. M. The electronic book. *In Suarez, Michael Felix, and H. R. Woudhuysen. The Oxford Companion to the Book. Oxford: Oxford University Press*, page 164, 2010.

[22] R. Grasset, A. Dunser, and M. Billinghurst. The design of a mixed-reality book: Is it still a real book? In *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*, pages 99 –102, 2008.

[23] S. Gupta and C. Jaynes. The universal media book: tracking and augmenting moving surfaces with projected information. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 177–180, Washington, DC, USA, 2006. IEEE Computer Society.

[24] A. Haans and W. IJsselsteijn. Mediated social touch: a review of current research and future directions. *Virtual Real.*, 9(2):149–159, 2006.

[25] B. L. Harrison. E-books and the future of reading. *IEEE Comput. Graph. Appl.*, 20:32–39, May 2000.

[26] E. Hoggan, S. A. Brewster, and J. Johnston. Investigating the effectiveness of tactile feedback for mobile touchscreens. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1573–1582, New York, NY, USA, 2008. ACM.

[27] L. Hong, Ed H. Chi, and S. K. Card. Annotating 3d electronic books. In *CHI '05 extended abstracts on Human factors in computing systems*, CHI '05, pages 1463–1466, USA, 2005. ACM.

[28] C. E. Izard. Human emotions. *New York: Plenum Press*, 1977.

[29] H. Kato. Development of artoolkit for augmented reality systemb. In *PRMU 2001(in Japanese)*, pages 79–86, 2002.

[30] T. Kristina and D. M. Christopher. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *In EMNLP/VLC 2000*, pages 63–70, 2000.

[31] M. Lebert. A short history of ebooks. *Retrieved March*, 2:2010, 2009.

[32] Y. Lee, T. Ha, H. Lee, K. Kim, and W. Woo. Digilog book-convergence of analog book and digital content. *Korea Information Processing Society*, 14:186–189, 2007.

[33] P. Lemmens, F. Crompvoets, D. Brokken, J. van den Eerenbeemd, and G.-J. de Vries. A body-conforming tactile jacket to enrich movie viewing. In *EuroHaptics*

conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint, pages 7 –12, march 2009.

[34] V.I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet Physics Doklady*, volume 10, pages 707–710, 1966.

[35] R. Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 140:1–55, 1932.

[36] Netlingo List. Netlingo list of chat acronyms @ONLINE. May 2010.

[37] B. Magnini and G. Cavagli. Integrating subject eld codes into wordnet. *In Proc. of the Second International Conference on Language Re-sources and Evaluation. Athens, Greece*, pages 1413–1418, 2010.

[38] A. Mangent. Hypertext fiction reading: haptics and immersion. *Research in Reading*, 31(4):404 –419, nov. 2008.

[39] H.J. Moll-Carrillo, G. Salomon, M. Marsh, J.F. Suri, and P. Spreenberg. Articulating a metaphor through user-centered design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 566–572. ACM Press/Addison-Wesley Publishing Co., 1995.

[40] M. J. Morón, R. Luque, E. Casilari, and A. Díaz-Estrella. An analytical model for estimating the delay in bluetooth communications with serial port profile. In *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, IWCMC '09, pages 1178–1183, New York, NY, USA, 2009. ACM.

[41] M. Mulder, A. Nijholt, M. Den Uyl, and P. Terpstra. A lexical grammatical implementation of affect. In *Text, Speech and Dialogue*, pages 171–177. Springer, 2004.

[42] A. Neviarouskaya, H. Prendinger, and M. Ishizuka. Emoheart: Conveying emotions in second life based on affect sensing from text. *Advances in Human-Computer Interaction*, 2010.

[43] A. Ortony and J. J. Turner. What's basic about basic emotions? *Psychological Review*, pages 315–331, 1990.

[44] A. Park, Kyoung J. Lee, and F. Casalegno. The three dimensions of book evolution in ubiquitous computing age: Digitalization, augmentation, and hypermediation. In *Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), 2010 IEEE International Conference on*, pages 374 –378, june 2010.

[45] J. Park and W. Woo. Multi-layer based authoring tool for digilog book. In Stphane Natkin and Jrme Dupire, editors, *Entertainment Computing ICEC 2009*, volume 5709 of *Lecture Notes in Computer Science*, pages 234–239. Springer Berlin / Heidelberg, 2009. 10.1007/978-3-642-04052-8_26.

[46] T. Park, J. Hwang, and W. Hwang. Facilitating the design of vibration for handheld devices. In Julie Jacko, editor, *Human-Computer Interaction. Novel Interaction Methods and Techniques*, volume 5611 of *Lecture Notes in Computer Science*, pages 496–502. Springer Berlin / Heidelberg, 2009.

[47] W. Parrott. Emotions in social psychology. *Psychology Press, Philadelphia*, 2001.

[48] W.G. Parrott. *Emotions in social psychology: essential readings*. Psychology Pr, 2001.

[49] Phelios. Phelios @ONLINE. May 2010.

[50] B. Pietro, O. Roberto, and H. Blake. Architectures for shared haptic virtual environments. *Computers & Graphics*, 21(4):421 – 429, 1997. Haptic Displays in Virtual Environments and Computer Graphics in Korea.

[51] R. Plutchik. A general psychoevolutionary theory of emotion. *In R. Plutchik & H. Kellerman (Eds.), Emotion: Theory, research, and experience: Vol. 1. Theories of emotion, New York: Academic*, pages 3–33, 1980.

[52] S. Poslad. *Ubiquitous Computing*. Wiley Online Library, 2009.

[53] A. S. M. M. Rahman, S. A. Hossain, and A. El Saddik. Bridging the gap between virtual and real world by bringing an interpersonal haptic communication system in second life. In *IEEE International Symposium on Multimedia*, Taiwan, Nov 2010.

[54] S. Réhman and L. Liu. ifeeling: Vibrotactile rendering of human emotions on mobile phones. *Mobile Multimedia Processing*, pages 1–20, 2010.

[55] English Grammar Revolution. List of interjections @ONLINE. May 2010.

[56] V. Ro. Wordnet-affect: an affective extension of wordnet. In *In Proceedings of the 4th International Conference on Language Resources and Evaluation*, pages 1083–1086, 2004.

[57] B.N. Schilit, M.N. Price, G. Golovchinsky, K. Tanaka, and C.C. Marshall. The reading appliance revolution. *Computer*, 32(1):65 –73, jan. 1999.

[58] K. Seiji and I. Yutaka. Perception of synchronization errors in haptic and visual communications. *Proc. SPIE 6391, 639108*, 2006.

[59] J. D. Steven and V. D. Andries. Document structure and markup in the fress hypertext system. *Markup Languages 1(1)*, pages 7–32.

[60] Theodor H. N. Steven C., Walter G., R. David, and V. D. Andries. A hypertext editing system for the 360. *Pertinent Concepts in Computer Graphics: Proceedings of the Second 17 University of Illinois Conference on Computer Graphics in Faiman and Nievergelt (eds.)*, pages 291–330, 1969.

[61] C. Strapparava and A. Valitutti. Wordnet-affect: an affective extension of wordnet. In *Proceedings of LREC*, volume 4, pages 1083–1086. Citeseer, 2004.

[62] N. Taketa, K. Hayashi, H. Kato, and S. Noshida. Virtual pop-up book based on augmented reality. In Michael Smith and Gavriel Salvendy, editors, *Human Interface and the Management of Information. Interacting in Information Environments*, volume 4558 of *Lecture Notes in Computer Science*, pages 475–484. Springer Berlin / Heidelberg, 2007. 10.1007/978-3-540-73354-6_52.

[63] Techdictionary. Techdictionary.com @ONLINE. May 2010.

[64] S.S. Tomkins. Affect theory. *In K. R. Scherer & P. Ekman (Eds.), Approaches to emotion, Lawrence Erlbaum, Hillsdale, NJ*, pages 163–196, 1984.

[65] D. Tsetserukou, A. Neviarouskaya, H. Prendinger, N. Kawakami, and S. Tachi. Affective haptics in emotional communication. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–6. IEEE, 2009.

[66] T. Vaughan. Multimedia make it work. *Mc Graw Hill*, 2008.

[67] Webopedia. Webopedia @ONLINE. May 2010.

[68] G. Welch, A. State, A. Ilie, L. Kok-Lim, A. Lastra, B. Cairns, H. Towles, H. Fuchs, Y. Ruigang, S. Becker, D. Russo, J. Funaro, and A. van Dam. Immersive electronic books for surgical training. *Multimedia, IEEE*, 12(3):22 –35, 2005.

[69] wikipedia. Java platform, micro edition @ONLINE. 2011.

[70] wikipedia. Java platform, standard edition@ONLINE. 2011.

[71] wikipedia. Portable document format@ONLINE. 2011.

[72] Windweaver. Windweaver emoticon liste @ONLINE. May 2010.

[73] C.H. Wu, Z.J. Chuang, and Y.C. Lin. Emotion recognition from text using semantic labels and separable mixture models. *ACM Transactions on Asian Language Information Processing (TALIP)*, 5(2):165–183, 2006.

Appendix A

Modified Cinderella Fairy Tale

Once upon a time there was a beautiful girl called Cinderella. She had two ugly step sisters. They were very unkind and made her do all the hard work. She had to sweep the floors, do all the dishes etc. At many time, they dress up in fancy clothes and go to lots of parties. One day a special invitation arrived at Cinderella's house from the royal palace which is ***50 miles away***. The king's only son was a truly handsome prince was going to have a grand ball. ***Three girls*** were invited to come. Cinderella knew she wouldn't be allowed to go to the ball. But the ugly sisters, ho ho ho, they were excited. They couldn't talk about anything else. When the day of the ball came, they made such a fuss. Poor Cinderella had to rush about upstairs and downstairs. She fixed their hair in fancy waves and curls. She helped them put on their expensive new ***white and pink*** dresses. And she arranged their ***golden*** jewels just so.

As soon as they had gone, Cinderella sat down by the ***black*** fire and she said. "Oh I do wish I could go to the ball". The next moment, standing beside her was a lovely old lady with a ***gold*** wand in here hand. "Cinderella, she said "I am your fairy godmother and you shall go to the ball. But first you must go into the garden and pick a red pumpkin, then bring me six mice from the mousetraps, a whiskered rat from the rat trap, and six lizards. You'll find the lizards behind the watering can. So Cinderella fetched a golden pumpkin, six grey mice, a whiskered rate, ***six lizards***. The fairy godmother touched them with her wand and the pumpkin became a golden coach, the mice became six grey horses, the rat became a coachman with the most enormous moustache, and the lizards became six footmen dressed in green and yellow, then the fairy godmother touched Cinderella with the wand and her old dress became a ***white*** dress sparkling with jewels while on her feet was the prettiest pair of ***golden*** slippers ever seen.

Remember said the fairy godmother you must leave the ball before the clock strikes ***two*** because at midnight the magic ends. "Thank you fairy godmother" said Cinderella and she climbed into the coach. The coach travelled the distance in ***17 minutes***. When Cinderella arrived at the ball she looked so beautiful that everyone wondered who she was! Even the ugly sisters. The Prince of course asked here to dance with him and they danced all evening. He would not dance with anyone else. Now Cinderella was enjoying the ball so much that she forgot her fairy godmothers warning until it was almost midnight and the clock began to strike.

One. Two. Three. She hurried out of the ballroom. Four. Five. Six. As she ran down the palace steps one of her glass slippers fell off on ***3rd stair***. Seven. Eight. Nine. She ran on toward the golden coach. Ten Eleven Twelve. Then there was Cinderella in her old dress. The golden pumpkin lay in her feet. And scampering down off down the road were six grey mice, a whiskered rat and six green lizards. So Cinderella had to ***run home*** and by the time the ugly sisters returned home was sitting quietly by the ***white fire***. Now when Cinderella ran from the palace, the prince tried to follow her and he found the glass slipper. He said, "I shall marry the beautiful girl whose foot fits this slipper and only her. In the morning the prince went from house to house with the glass slipper and every young lady tried to squeeze her foot into it. But it didn't' fit any of them.

At last the prince came to Cinderella's house on horse. First one ugly sister tried to squash her foot into the ***size 6 slipper***. But her foot was too wide and fat. Then the other ugly sister tried but her foot was too long and thin. Please said Cinderella, let me try. "The slipper won't fit you", said the ugly sisters. "You didn't go to the ball!" But Cinderella slipped her foot into the glass slipper and it fit perfectly. The next moment standing beside her was the fairy godmother. She touched Cinderella with the wand and there she was in a golden dress sparkling with jewels and on her feet was the prettiest pair of glass slippers ever seen. The ugly sisters were so surprised that, for once they couldn't think of anything to say. But the ***black coated Prince*** knew what to say. He asked Cinderella to marry him. And then there was a happy wedding. Everyone who had gone to the ball was invited, even the ugly sisters. There was wonderful food, lots of music and dancing. And the Prince of course danced every dance with Cinderella. He would not dance with anyone else