Digilog book for temple bell tolling experience based on interactive augmented reality

Taejin Ha · Youngho Lee · Woontack Woo

Abstract We first present the concept of the Digilog Book, an augmented paper book that provides additional multimedia content stimulating readers' five senses using augmented reality (AR) technologies. We also develop a prototype to show the usefulness and effectiveness of the book. The Digilog book has the following characteristics: AR content descriptions for updatable multisensory AR contents; enhanced experience with multisensory feedback; and interactive experience with computerized vision-based manual input methods. As an example of an entertaining and interactive Digilog Book, this paper presents a "temple bell experience" book and its implementation details. Informal user observation and interviews were conducted to verify the feasibility of the prototype book. As a result, this case study of the Digilog book can be useful in guiding the design and implementation of other Digilog applications, including posters, pictures, newspapers, and sign boards.

Keywords Digilog book · Culture technology · User interaction · Multisensory experience · Augmented reality

T. Ha · W. Woo
GIST U-VR Lab, Gwangju 500-712, South Korea
e-mail: tha@gist.ac.kr

W. Woo
e-mail: wwoo@gist.ac.kr

Y. Lee (✉)
MNU U-VR Lab, 61 Dorim-ri, Cheonggye-myeon,
Muan-gun, Jeonnam, Korea
e-mail: youngho@mokpo.ac.kr

1 Introduction

Electronic books (e-books) have been made available through computers, PDAs, and even mobile phones. The first e-books were produced in text formats (e.g., PDF, XML), while the second-generation e-books combined multimedia content. Despite early expectations that e-books would replace paper books, this has not been the case (Sellen and Harper 2003). There are several reasons that people still prefer paper books: physical presence (tangibility), possession, and the high quality of printed material. On the other hand, digital content has other advantages, including portability and additional information like multisensory feedback. To this end, augmented reality (AR) books integrate the advantages of paper books with digital content, so users can experience both analog aesthetic emotions and immersive digital multisensory feedback.

Many researchers have examined AR books, beginning with The MagicBook (Billinghurst et al. 2001) that demonstrate the potential and usability of AR books as new-generation media. After the work was published, several implementations of AR books were created for education, storytelling, simulation, game, and artwork purposes. The most supported interactions in the AR book were viewing and listening with turning pages. In other works, various interactions involved 3D magnetic trackers, foldable cube, paddle, gaze tracking, mouse click, and turntable with a slide device.

This paper proposes Digilog Book, a next-generation publication that combines the analog sensibility of a paper book with digitized visual, auditory, and haptic feedback by exploiting AR technology. The Digilog Book consists of a conventional printed book, multimedia content, and a Digilog Book viewer that acquires images of the printed book from the camera and augments the multimedia content in the book. This allows the reader to experience the benefits of printed

material enhanced by multimedia content through multisensory visual, auditory, and haptic experiences. In other words, a Digilog Book maintains the original functions of a printed book while adding the advantages of digital content.

The Digilog book has the following characteristics: AR content descriptions for updatable multisensory AR contents; enhanced experience with multisensory feedback; and interactive experience with computerized, vision-based, bare-hand input methods.

The publication of previously implemented digital content would be difficult to update given the changing requirements of conventional AR book applications. To solve this problem, this paper proposes AR content descriptions in which the AR content of the Digilog Book will be updated through an Internet connection if AR content in a remote database is updated. Moreover, this project adds vibration feedback to visual and audio feedbacks via a 3D manipulation tool in which a vibration module is embedded. This can create an immersive experience with visual, audio, and enhanced haptic feedback that is possible with a paper book. This project also proposes an input method for natural interaction with AR content, requiring only the user's hands. In an early occlusion-based virtual button, the user manually occludes a fiducial marker on a page in order to trigger an event. This paper suggests an improved method in which unnatural fiducial markers do not need to be inserted on the book pages. Three particular types of virtual buttons are proposed that support various input types with respect to interaction scenarios.

As an example of an entertaining and interactive Digilog Book, this paper presents a "temple bell experience" book that explains Asian cultural heritage to users in a way that a conventional book cannot. For example, users can rotate and observe a virtual temple bell by covering virtual buttons on the book with their fingers or by tapping the 3D object with a manipulation tool; corresponding visuals, sounds, and vibration feedback are then initiated. Sequential start, emersion/withdrawal, interaction, and the transitional interpolation of the 3D object and visual special effects are also considered for natural augmentation. Through these interactions, the temple bell experience book is expected to encourage readers to explore other cultural heritages for education and entertainment.

In considering a practical running environment for typical users, this project designed a monitor-based display installation with a camera fixed on a table in line with the direction of a user for natural viewing. The typical USB camera used in this project is less expensive than a high-end camera, and the manipulation tool is simply a modification and extension of a conventional mouse. Therefore, this is an appropriate choice for home users. In order to verify the feasibility of implementation, we conducted informal user observation and interviews.

The remainder of this paper is organized as follows: Sect. 2 explains the related work and characteristics of the Digilog Book. Section 3 describes the system design that enables AR interaction with multisensory feedback. As an example, Sect. 4 shows the implementation details and results for an AR-based temple bell experience book. Finally, lessons and future directions of the Digilog Book are summarized in Sect. 5.

2 Related work

2.1 Cultural technology in AR

Over the past few decades, the computing paradigm has been changing from mainframe computing, network computing, and personal computing to social computing, or, with respect to the viewpoint of the user, user-centric computing. From such paradigm shifts, cultural technology (CT) has emerged, drawing a great deal of attention to cultural aspects in the social computing paradigm. In general, CT is, from a narrow point of view, the digitalizing of cultural contents through technology and, from a wider point of view, technology's improving quality of life for humans by creating culture. CT-related industries are wide in range: film, broadcasting, gaming, music, art, cultural heritage, tourism, and computing (Tosa et al. 2005; Rauterberg 2006).

Several AR-related projects have been launched to enhance humans' imagination, creativity, and sensitivity with CT. In the Mobile Augmented Reality System (MARS) project (Höllerer et al. 1999), users can experience images, video, narration, and 3D models related to a certain building on campus explaining its historical significance. The Archeoguide project provides an interactive AR guide for archaeological sites (Stricker et al. 2001). In this project, users use a wearable AR system to visualize the Olympics 2,000 years ago and to understand significant events in Grecian heritage. In museum applications, 3D models of artifacts are created and rendered in an AR environment through photo-realistic and efficient 3D modeling (White et al. 2003). Viewing text, images, and 3D models in the same AR environment enhances the user's experience. Additionally, in traditional architecture applications, users can understand the traditional structure and culture of the "Young's Ancestral Hall" in Taiwan using AR technology (Chen et al. 2009).

2.2 Related work for Digilog book

Many researchers have examined AR books, beginning with "The MagicBook" (Billinghurst et al. 2001), which demonstrates the potential and usability of AR books as

new-generation media. In this application, a user holding a handheld display that contains a small camera can experience the full reality of a virtual continuum through a paper book. Since Billinghurst's work was published, other AR books have been created and proposed to teach details of earth-sun relationships in undergraduate geography courses (Shelton and Hedley 2002). In this system, viewing interaction uses the head-mounted display (HMD), which holds a USB camera. In educational software for Kanji Learning (Wagner and Barakonyi 2003), multiple users holding PDA viewers can play a card game in a collaborative AR environment. Saso et al. propose an AR storytelling book, "Little Red," which illustrates the story of Little Red Riding Hood (Saso et al. 2003). Users can partly control their interactions with the book by using their hands and wearing HMDs with USB cameras. The user can change the story and its ending. In the Vivid Encyclopedia's "MR Pictorial Book of Insects" (Shibata et al. 2004), the user can interact with virtual insects using a 3D magnetic tracker. "EyeMagic Book" (McKenzie and Darnell 2004), a storytelling book application, uses augmented reality storytelling in the context of a children's workshop. A user in front of the application's kiosk can view a 3D object and various animations using a handheld display that has a small camera. In the "AR Volcano Kiosk" (Woods et al. 2004), a user can interact with augmented 3D objects through a turntable with a slider while holding a display that has a small camera. "Little Feet and Big Feet" and "Looking for the Sun," by Dünser et al. (2007), are AR storybook experiences created for young children. In this system, a web-cam is mounted on top of a computer screen. As such, it can be used in most modern classrooms or homes. For interaction, users use a paddle to trigger certain events and a mouse to click on navigation buttons that advance to the next or previous text page and allow the child to listen to the story.

Another AR book by Grasset et al. (2008) is based on "The House that Jack Built;" with this storytelling book, a user can experience 2D and 3D animation and 3D sound through a handheld display and a small camera. Paddle interaction and gaze interaction are supported. Zhou et al. (2008) present wIzQubes™, a novel user interface that supports interactive storytelling in an AR environment. It enhances the corporeal interactions of traditional AR books by using a foldable cube to "turn" the "page" in a monitor-based display system with different viewpoints between the user and the camera. Two cubes are used to navigate through various scenes of the story and choose different items that are needed in the story. The Eye of Judgment is an AR card game based on a monitor-based display that presents various viewpoints between the user and the camera (The Eye of Judgment). Interaction is conducted by positioning cards in a layout with proximity of distance. The

"Virtual pop-up book" by Taketa et al. (2007) provides two kinds of display methods. The PC monitor display type is best suited for use by children who find HMD too cumbersome or for demonstrations in which many users see the content simultaneously. On the other hand, the HMD type gives the user a sense of immersion. Viewing with turning-page interaction is also supported. "The Haunted Book," by Scherrer et al. (2008), is an interesting AR artwork application that is based on a poem. It features 2D animated illustration augmentation instead of 3D objects. In this installation, a computer screen is used to avoid the restrictions of wearing the HMD. This system also provides simple viewing (with turning pages) interaction. Cho et al. (2007) and Yang et al. (2008) have suggested creating an e-learning system application through AR books, targeting elementary school students. Viewing with turning pages is a supported interaction in the project.

2.3 Characteristics of Digilog book

2.3.1 Features of the AR-based experience environment

AR visually provides additional and meaningful virtual content (information) about a practically observed object in a current situation. On the other hand, in a virtual reality (VR)-based experience environment, all scenes are represented to the user as virtual objects in a specific computing space (e.g., CAVE or VR simulation environment). Therefore, AR can support more enhanced immersion and realism in real space because the input images used as background in an AR environment can achieve a more realistic and familiar effect than in VR. Additionally, a user can directly manipulate augmented information about a real object in a real environment in order to gain a more in-depth understanding of the real object. Many advantages of an AR-based experience system have been reported (Azuma 1997).

Based on the literature, this project has developed four characteristics for the Digilog book.

2.3.2 AR content descriptions for updatable, multisensory AR content

Content-sharing or collaborative content generation by Internet users has become common. Previously implemented virtual content publication, however, would be difficult to update in conventional AR book applications. If content modification occurred, then additional modeling and programing works would be inevitable.

To solve this problem, this project proposes a file format for AR content descriptions in which the AR content of the Digilog Book will be updatable through an Internet connection if AR content in a remote database is updated. The

suggested file format also includes several properties for visual, audio, and haptic feedback. This is necessary for an immersive interactive experience with multisensory feedback.

2.3.3 Enhanced multisensory feedback to support an immersive experience

Most of the feedback in the conventional AR book application is visual (e.g., seeing the book), spoken (e.g., narration), and tactile (e.g., turning pages). This paper adds vibration feedback via a vibration module embedded in the proposed 3D manipulation tool. This can motivate an immersive experience through visual, audio, and enhanced haptic feedback, which conventional paper books cannot provide.

2.3.4 A natural, computerized, vision-based bare-hand input method for an interactive experience

This paper proposes an input method for natural interaction with AR content that requires only the user's bare hands. In an early iteration of an occlusion-based virtual button, the user occludes a fiducial marker with his or her hands in order to trigger an event (Lee et al. 2004). This project suggests an improved method in which unnatural markers need not be inserted on the book pages and proposes three additional types of virtual buttons that support various input types with respect to interaction characteristics.

2.3.5 Concrete interaction design

The target users for this project are typical users with home PCs. While considering a practical running environment, a monitor-based display installation was designed with a USB camera fixed on a table in line with the direction of a user for natural ergonomic viewing. The monitor-based display system used is less expensive and more popular than a head-mounted display (HMD) system, and the designed manipulation tool is simply a modification and extension of a conventional mouse input device. Therefore, this is the proper choice for home users in their normal computer environments.

In order to verify the feasibility of implementation, this project conducted informal user observation and interviews. Lessons and future directions of the Digilog Book are summarized in the discussion section.

3 Digilog book

3.1 System overview

Figure 1 shows the Digilog Book system architecture based on the proposed interactive AR system. Based on input images from a camera, a computer vision-based tracking manager recognizes and tracks a paper book, a manipulation tool, and a hand object. The collision detector then inspects penetrations between a virtual line created by the
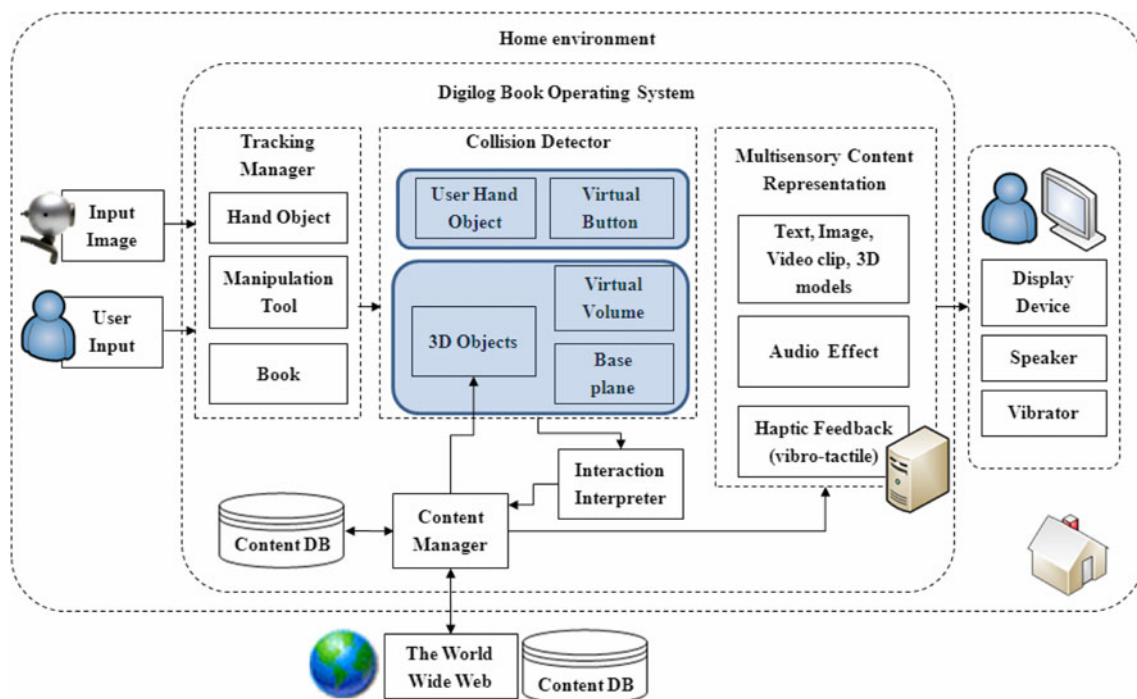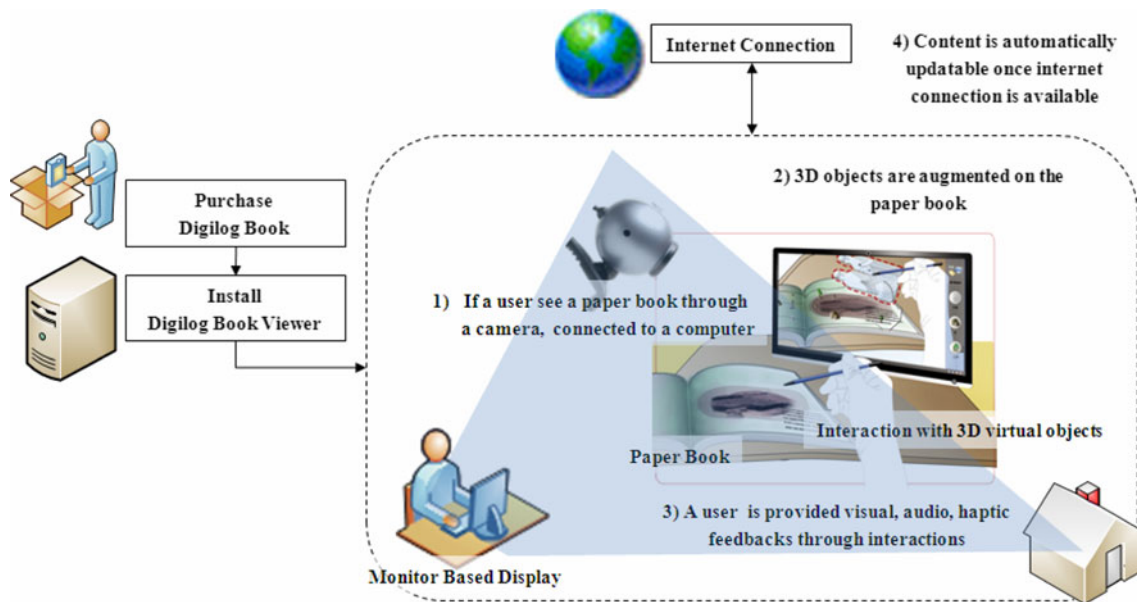

Fig. 1 Block diagram of Digilog Book viewer

Fig. 2 Conceptual figure of Digilog Book usage

manipulation tool and a bounding volume of the augmented 3D objects that are based on the book. The detector also checks an occluded area between the virtual buttons and the user's hand objects. At this point, the interaction interpreter conducts examinations like 3D object pointing and hitting, movement interactions, or hand interactions for pushing virtual buttons.

Next, the content manager composes proper multisensory content in order to react to the user input. The multisensory content consists of visual feedback (text, images, video clips, and 3D models), audio feedback (effects and background sounds), and haptic feedback (vibro-tactile interactions). Finally, a display device, a speaker, and a vibrator of the manipulation tool represent the multisensory content. Additionally, if AR content in a remote database is updated, then the AR content of the Digilog Book will be updatable through an Internet connection.

Figure 2 shows a conceptual flow of Digilog Book usage, from purchase to interactive AR experience. A user purchases a Digilog Book and installs a Digilog Book viewer on a home computer. While the viewer is running, based on captured images from a camera, the computer detects a specific fiducial of the book, computes the camera's coordinate relation from the book, and sets the basis coordinate. The 3D models, video clips, sounds, or other content are augmented based on the basis coordinate. While tracking the fiducial, the content is still augmented in the book. The user interacts with the augmented virtual content using a manipulation tool or virtual buttons, and he or she can simultaneously see, hear, and physically sense the augmented content from the paper book. The AR content is updatable if an Internet connection is available.

Fig. 3 Running environment of Digilog Book viewer; the monitor, camera, and book were aligned

In this real demonstration, the setting, monitor, camera, and book were aligned, as shown in Fig. 3 and were all in the user's direct view. The camera view faced the book so that the entire book area was contained within the camera's viewing angle.

3.1.1 Target user/content/interaction

The target user is an average computer user who has no related AR, computer vision- or graphics-related knowledge, or programing skills. Specifically, the target reader is a child in the elementary-school age range. As an example of a Digilog Book, this project proposes a temple bell experience book that explains Asian cultural heritage in a way that provides multisensory feedback and natural interaction. For example, users can translate or rotate a virtual temple bell, point to facets of the temple bell to play multimedia documents, or hear a tolling bell and sense

vibration feedback by tapping the 3D bell model. This type of immersive experience can motivate readers to seek cultural heritages for education and entertainment purposes. For the user interface to be easy and intuitive, the user should be able to understand its potential simply by adopting familiar metaphors for interaction without resorting to a manual.

3.1.2 Environment/hardware

The Digilog Book viewer is executed in an arbitrary indoor environment with sufficient lighting conditions to recognize the computer vision-based fiducial pattern. The only special requirements for the program are a desk (for laying out a book), a camera (for acquiring book images), and a computer (for processing the input images). An ordinary low-cost USB camera is fixed on a camera arm that is attached to the desk. This monitor-based AR display installation is suitable because it can be used in most home environments without any expensive additional hardware.

A wearable display type (e.g., optical see-through, video see-through) has several disadvantages; its cost is relatively high compared to that of the monitor-based AR display, and wearing the HMD or holding the device may cause arm or wrist fatigue or even dizziness. Instead, because the camera is moving, 3D object observation is possible from arbitrary positions. Tracking errors, however, can be increased as a result of camera movement.

For tracking the book and the manipulation tool, this project uses fiducial pattern markers [ARToolKit]. Although this method produces limited accuracy compared to an infrared optical or magnetic tracking method, its cost is much lower and it is easier for an average user to use.

3.2 System design enabling AR interaction with multisensory feedback

3.2.1 AR content descriptions for updatable multisensory AR content

This paper proposes an AR content description that will update previously implemented virtual content from AR book publication in response to users' changing requirements. Therefore, the AR content of a Digilog Book could be updatable through an Internet connection if AR content in a remote database is updated. The suggested file format also includes several properties for visual, audio, and haptic feedbacks. This is necessary for an immersive, interactive experience with multisensory feedback.

Previous works suggest XML-based data descriptions or customized markup language to construct an AR scene

(Irawati et al. 2008; Lee and Kim 2009). However, there is a critical limitation on extension and compatibility with existing international standards. To solve this problem, this project exploits Extensible 3D (X3D). X3D is a software standard for describing an interactive virtual environment, and it is the successor of the Virtual Reality Modeling Language (VRML). X3D features 2D/3D graphics, animation, spatialized audio and video, user interaction, navigation, user-defined objects, scripting, networking, physical simulation, CAD geometry, programmable shaders, and particle systems. Among these functions, networking can compose a single X3D scene from assets located on a network. By setting a Uniform Resource Locator (URL) field and an activating condition, remote content can be loaded.

Visual descriptions of a Digilog Book could contain properties such as translation, rotation, scaling of the augmented 3D model, text, figures, video clips, and even color, texture, and lighting. For audio feedback, the Digilog Book should support auditory descriptions, such as effects, background sound, volume, sound direction, and tempo, in response to user interaction. 3D audio effects may be required for a more realistic or impressive experience. Most visual and audio effects can be described by functions of X3D support. Haptic feedback properties are not currently supported by X3D standard, but they are possible with the extended description of Digilog Book. The considered vibration modules are embedded internally in the manipulation tool. Vibration modules are activated when a user causes a collision between some virtual objects. In the Digilog Book, the vibration module of the manipulation tool can be activated depending on the collision direction and speed with respect to a 3D object. The intensity, direction, and duration time of each vibration should be described.

The Digilog Book could also include other properties, including a state and action trigger, events that occur with respect to distance between a 3D object and the manipulation prop, and collision detection so that when the user pushes the virtual buttons of the book, the corresponding reaction is activated.

By maintaining compatibility with existing X3D, the proposed description can be interpreted through the X3D viewer with limited functions. It can be viewed with full functions by adopting the proposed namespace, Extensible Markup Language (XML) schema definition (XSD), and implementation modules. Table 1 shows extended nodes and fields of the Digilog Book that are not contained in existing X3D. The X3D can be represented by a variety of encoding formats, including XML. The implementation section shows an XML example for an AR scene using the extended description.

Table 1 Field definitions (specifications) for Digilog book data format (following the conventions of the ISO/IEC 19775)

Field type	Access type	Name	Default	Range
Visual/audio feedback node				
MFString	inputOutput	3Dmodel	NULL	[URL]
MFNode	inputOutput	Text	NULL	[Text]
MFNode	inputOutput	Audio	NULL	[AudioClip]
MFNode	inputOutput	Image	NULL	[ImageTexture]
MFNode	inputOutput	Video	NULL	[MovieTexture]
Haptic feedback (vibration) node				
SFString	inputOutput	Description	NULL	–
SFBool	inputOutput	Loop	False	–
SFNode	inputOutput	Metadata	NULL	[X3DMetadataObject]
SFFloat	inputOutput	Intensity	0	[0, 255]
SFTime	inputOutput	Duration	0	(0, ∞)
MFString	Output	url	NULL	[URL]
SFBool	Output	isActive	–	–
SFBool	Output	isPaused	–	–
Digilog book (for updatable contents) node				
SFBool	inputOutput	Load	True	–
SFNode	inputOutput	Metadata	NULL	[X3DMetadataObject]
MFString	inputOutput	url	NULL	[URL]

3.2.2 Design of a manipulation tool that provides vibration feedback with respect to interactions

For the average non-professional user in a normal computer environment, this project uses a cubical box with a virtual line that is attachable to a mouse input device. The mouse can easily be exploited in a normal computer environment. By using the mouse input buttons, a discrete input generates the event functions that are necessary for Digilog Book interaction. In addition, multiple fiducial patterns are printed on the box, enabling the manipulation tool tracking at arbitrary camera viewing angles and simultaneously providing continuous input for modifying a 3D object's position and rotation properties (Ha and Woo 2010).

This paper also suggests a vibro-tactile, haptic user interface. A mobile vibration module is used to enhance immersion when using the Digilog Book. Considering its portable, lightweight design, the vibration module can be embedded in the manipulation tool, where it directly stimulates the user's skin by providing vibro-tactile feedback. The user can select several vibration patterns for clicking, dragging, translating, rotating, pointing, or tapping events. The patterns consist of vibration waves in various shapes, such as a square wave, a triangle wave, a sin wave, and other shapes of waves (Lee et al. 2009).

The users can obtain additional information about the specific elements of the 3D object by pointing to the corresponding parts with the manipulation tool. Visual
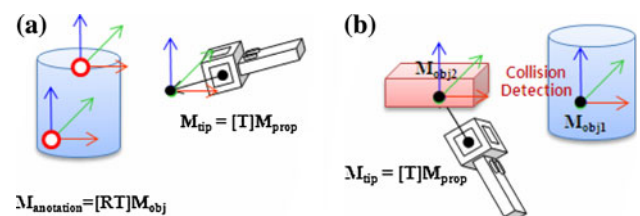
Fig. 4 Pointing and hitting interactions: **a** pointing method and **b** tapping method

annotations on various parts of the 3D object indicate a specific viewing point for figures, text, and video clips. Figure 4(a) shows the translated/rotated annotation points from the 3D object and an augmented line tip (M_{Tip}: pose matrix of the tip) on the front part of the prop (manipulation tool). If the Euclidian distance between these points is within a threshold distance, then a collision can occur. In Fig. 4, **R** is a rotation and **T** is a translation matrix.

To design the hitting interaction, a rigid body from the Open Dynamics Engine (ODE) is fitted onto the virtual object A (obj1) and onto another object B (obj2), and a real-time collision process is conducted between them. Using a drag-and-drop technique familiar to the user, 3 axis positions and 3 axis rotation manipulations (6DOF manipulation) of a 3D object are easily and simultaneously performed in the book page plane, even in a 3D space such as air. Thus, a user can intuitively select and move the 3D object. If the B object collides with the A object, then the Digilog Book viewer system provides the corresponding
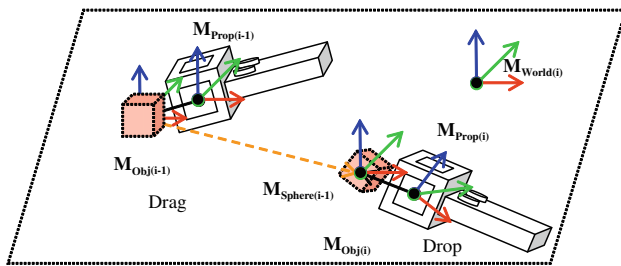
Fig. 5 3D object translation and rotation through drag and drop

visual, sound effect, and tactile vibration feedback to the user.

For the 6DOF manipulation, the collision detector first checks penetration depth between the virtual line at the front of the manipulation tool and a 3D object. Then, the user presses and holds a button for the manipulation tool to translate the 3D object to the wanted position. During the drag state, a 3D object position is obtained from the augmented sphere (M_{Sphere}) at the current frame, and rotation is set relative to the rotation of the sphere (M_{Sphere}) from the previous frame to the current frame. Equitation 1 shows these relationships, where R is a rotation and T is a translation matrix of the corresponding pose matrix (Fig. 5).

$$M_{Obj(i)} = R\big(M_{Obj(i-1)}\big) \times \Big[R\big(M_{Sphere(i-1)}\big)^{-1} R\big(M_{Sphere(i)}\big)\Big] T\big(M_{Sphere(i)}\big) \tag{1}$$

3.2.3 Virtual button interaction using a user's bare finger

Normally, virtual objects are vertically augmented on the base plane of a book; the book is on a table, and a fixed camera faces the book. In this arrangement, however, the view direction of the camera is from top to bottom; users observe the top and the forward (or backward) side of the augmented 3D object. Therefore, interaction methods are required to rotate the 3D objects as well as to trigger other interaction scenarios.

This paper adopts a virtual button interface that is natural, simple, and easy to control because it requires only the user's bare hands. In an early iteration of an occlusion-based virtual button, a user needed to occlude a fiducial marker with his or her hands in order to trigger an event (Lee et al. 2004). This project suggests an improved method in which unnatural markers need not be inserted into the book pages.

This method invisibly inserts virtual 3D buttons on the button area printed on the book. The buttons are projected
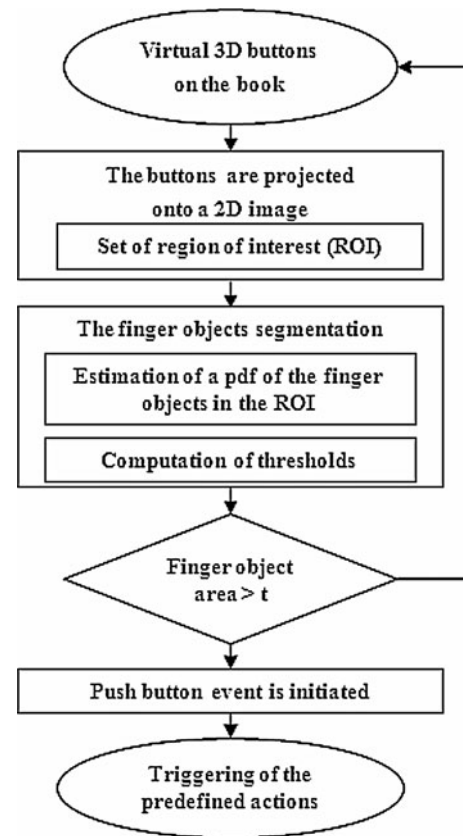

Fig. 6 Flow chart for virtual button interaction using a user's bare finger

onto a 2D image and are set by a region of interest (ROI). From this region, the finger objects segmentation process is run. Estimation of a color probability density function of the finger objects can then be conducted in real time to enable the stable segmentation of the finger objects under changing light conditions and to avoid the manual parameter settings (Ha et al. 2006). If the finger object area is over a specific threshold, then the push button event is initiated, which triggers reactions among the 3D objects (Fig. 6).

Three types of virtual buttons are also proposed to support various input types with respect to interaction characteristics. As shown in Table 2, three kinds of virtual buttons are considered for treating events in the Digilog Book: the toggle button, the push button, and the auto-releasable toggle button. Each button type is used in a particular situation. The toggle button changes the interaction mode, the push button is needed for an interaction that uses an instant time event, and the auto-releasable toggle button activates an interaction once the previous reaction completes (creating animation), similar to a synchronous event.

Table 2 Three types of virtual buttons and their functions

Button type	Function
Toggle button (mode change)	If the button is pushed once, then it remains in a pushed state until another button is pushed
Push button (asynchronous event)	If the button is not in a pushed state, then it is immediately released
Auto-releasable toggle button (synchronous event)	If the button is pushed once, then it remains in a pushed state until a corresponding event is triggered

Fig. 7 The Bell of King Seongdeok
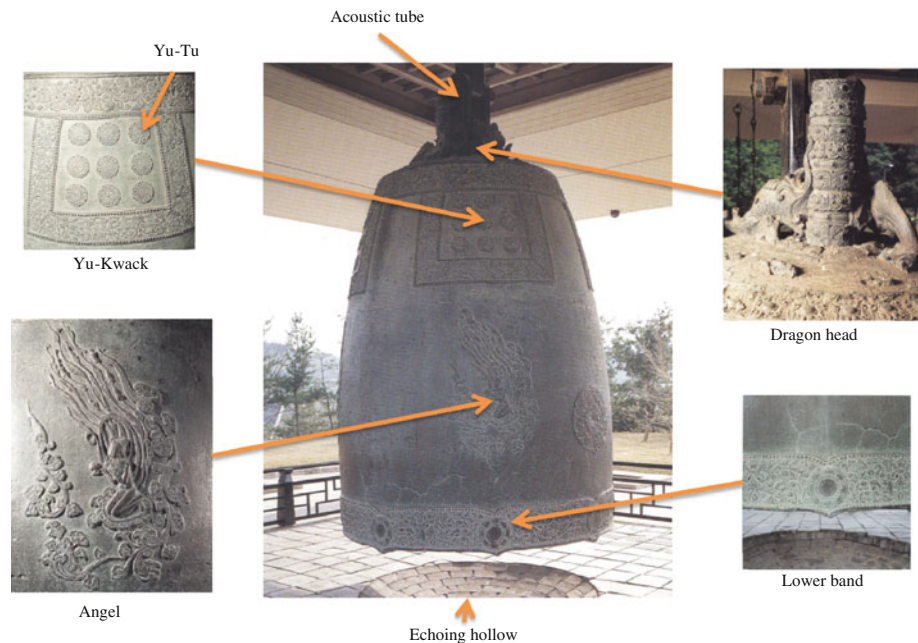
4 Temple bell experience book

4.1 Making the temple bell experience book

4.1.1 Concept of the temple bell experience book

The temple bell experience book is expected to encourage readers to explore cultural heritages for education and entertainment purposes. Figure 7 presents the patterns, shapes, qualities, and carvings of Korea's temple bells. The dragon head hangs on the bell tower, and the acoustic tube opposite the dragon head helps the sound of bell travel long distances. The lower band consists of repeated patterns. The echoing hollow is arranged to expand the reverberation with the least possible energy. The Yu-kwack consists of the Yu-kwack bands and the group of Yu-Tus, which resemble stalactites [Korean bells]. By using the Digilog book, readers can have enhanced multisensory and interactive experiences of the temple bells.

4.1.2 Temple bell content

In next step, multimedia content was created, including the Digilog Book, 3D models, 3D model mesh reductions, texture maps, sound effects, video clips, images, and other multimedia content. Figure 8 shows several 3D bell models that are used in this Temple Bell Experience Book. These scanned 3D models are supported by the documentation project for important movable cultural assets and by the cultural heritage administration of Korea [Cultural Heritage Administration of Korea].

4.1.3 Book design

Rectangular ARToolKit (2010) markers often cause disjointed relationships in a book's contents; this awkward visual display might diminish naturalism. From the book's contents, this project registers a related figure as a fiducial marker so that it is as realistic as possible among the

Fig. 8 Several 3D models of "The Temple Bell of Korea" Digilog Book

Fig. 9 Several pages of "The Temple Bell of Korea" book: **a** cover, **b** outline, and **c–d** temple bell contents

background text and figures. Figure 9 shows several of a total of 14 pages of the Temple Bell Experience Book. The left side shows text, a 4 × 4 cm marker, and figures, while the right side shows the temple bell figure and button images.

4.2 Implementation and results of interaction techniques of the temple bell experience book

This subchapter explains the implementation and results of the proposed interaction techniques. The program was

Fig. 10 An example of node syntax, a file format for interactive AR with multisensory feedback based on X3D standard

```
…
<X3D version="3.0" profile="Immersive">
  <head>
    <meta name="an example of a description file of Digilog Book" content="scene1.x3d"/>
  </head>
  <Scene>
    <Shape>
      # 3D model
      <3Dmodel DEF="BellModel" description="bell model" url="ARcontent/bell.x3d"/>
    </Shape>

<Sound>
  # audio feedback
  <AudioClip DEF="BellSound" description="bell sound" stopTime="10"
   url="ARcontent/bell.wav"/>
</Sound>
# haptic feedback
<Haptic DEF="BellHaptic" description="uniform vibration" intensity="128" duration="1"/>

# interaction node
<VirtualButtonInput DEF="VBI"/>
<ManipulationToolInput DEF="MTI"/>

# event flow
<ROUTE fromNode="VBI" fromField="touchTime" toNode="BellSound" toField="startTime"/>
<ROUTE fromNode="MTI" fromField="isCollision" toNode= "BellHaptic" toField="active"/>

# updatable AR contents
<DigilogBook DEF="NewARScene"
 url="http://www.DigilogBook/ARcontent/updatedContent.x3d" load="false"/>
  </Scene>
</X3D>
```

executed in a normal indoor environment with no dramatic changes in lighting. The camera was a general-purpose USB camera that was fixed on the camera arm and captured 30 frames per second at 640 × 480 pixel resolution. The computer was equipped with a 2.40 GHz CPU and 4 GB of memory. The osgART library (osgART) was used to support the rendering of scene-graph structured graphic models and the computer vision-based tracking functions.

4.2.1 AR content description with updatable multisensory feedback

This prototype provides an ideal platform for testing AR authoring that allows designers to vary the graphic, audio, and haptic feedback. One example of a description file based on X3D is represented in Fig. 10. In the Digilog Book node, if the load field is set to TRUE value, the X3D file described by the URL field can be loaded immediately.

4.2.2 Sequential steps of the start-interaction-end process

From the Digilog book viewer's initial state, the visual special effects are enhanced. The temple bell model is naturally augmented from the page of a book through vertical translation and horizontal rotation interpolation. Lastly, in the interaction step, the user can receive a detailed explanation or can manipulate the augmented temple bell model by rotating or hitting the bell.

As shown in Fig. 11, figures of books are 2D images, and augmentation objects are 3D models. For natural registration, geometrical consistency of appearance is required between the figures in the book and the 3D virtual objects. As a partial solution, this implementation interpolates rotations and translations between the 2D image figure and the augmented object. Specifically, the 3D temple bell model vertically rises from the 2D image at the same scale and then rotates frontward toward the user. During this interpolation process, the addition of a particle object enhances the lighting, making the visual effect more interesting and somewhat compensating for the inaccurate matching between the figures and the augmenting 3D models. This particle object is rendered through the alpha blending of a total of 150 texture sequences. In addition, the augmented 3D temple bell model repeatedly moves from left to right and vice versa, based on a vertical axis; this movement may compensate for the jittery visual effects caused by unstable marker tracking.

4.2.3 Manipulation tool that supports vibration feedback

A commercial pen mouse (wireless presenter) was adopted as a prototype for the implemented manipulation
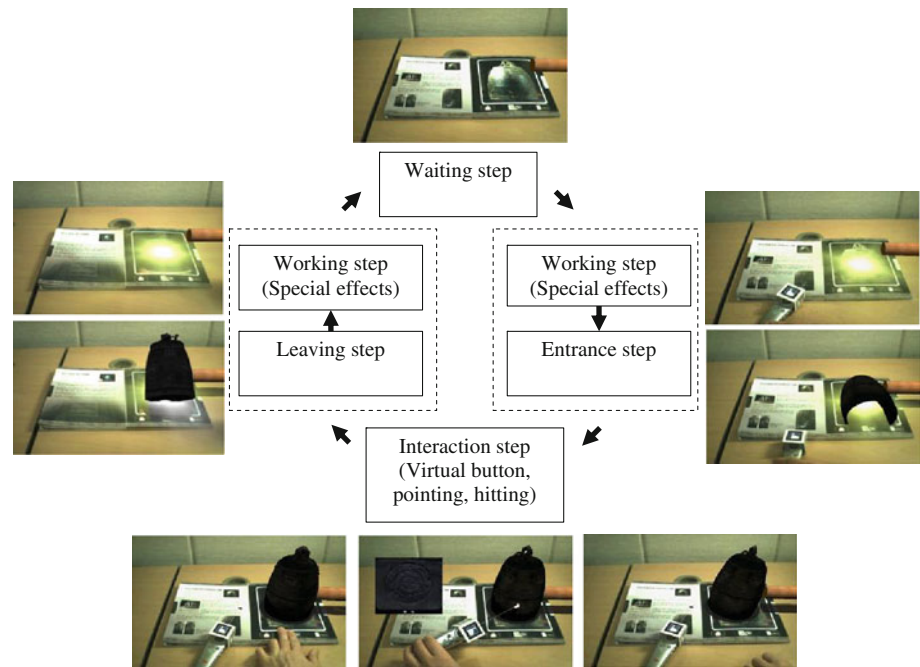
Fig. 12 The internal components of the manipulation tool included a vibration actuator, a Bluetooth communication module, and a microcontroller

tool. This was attached to a box with multiple markers fixed on its front. The marker size was 3 × 3 cm, which was determined by the length between the user's viewpoint and the recognizable marker when the user extended an arm while holding the manipulation tool. The internal components of the manipulation tool consisted of a vibration actuator, a Bluetooth communication module for wireless communication, and a microcontroller for tactile feedback generation and wireless communication (Fig. 12).

In Fig. 13(a), white spheres are augmented on specific parts of the temple bell model, and these provide visual annotations on the heads-up display (HUD) in order to indicate areas that contain information about the bell, including figures, texts, and video clips. By pointing to

them with a manipulation tool, users can obtain this additional information. In the event that a user hits the bell, as shown in Fig. 13(b), a cylindrical wooden model is selected and moved to the temple bell model with the manipulation tool. When the wooden model collides with the bell model, a bell sound is played.

4.2.4 Virtual button input triggered by the user's finger

Four 3D buttons are projected onto a 2D image, setting the ROI as shown in Fig. 14. If a finger covers more than 2/3 of the corresponding button area, then a push button event occurs. Computerized vision-based image processing is sequentially processed from the first button to the last button in an attempt to verify whether or not the user's hand is covering the printed button area of the book. If the system determines that the user is pressing a certain button and an evaluation of the remaining buttons can be omitted, then the image processing time can be shortened.

In the proposed interaction scenario, as shown in Fig. 15, the first button is a toggle-type button; once it is selected, it remains in a selected state until another button is selected. This button is used to change interaction mode, i.e., showing detailed explanations of specific parts of the temple bell with a manipulation tool. The second and third buttons are push-button types that trigger the rotation of the 3D model as long as the buttons are being pushed. The fourth button is an auto-releasable toggle button that is released as soon as the hitting bell event is triggered.

Fig. 13 Pointing and hitting interactions: **a** specified explanations obtained by pointing to specified parts of the temple bell; **b** using a manipulation tool to hit the temple bell
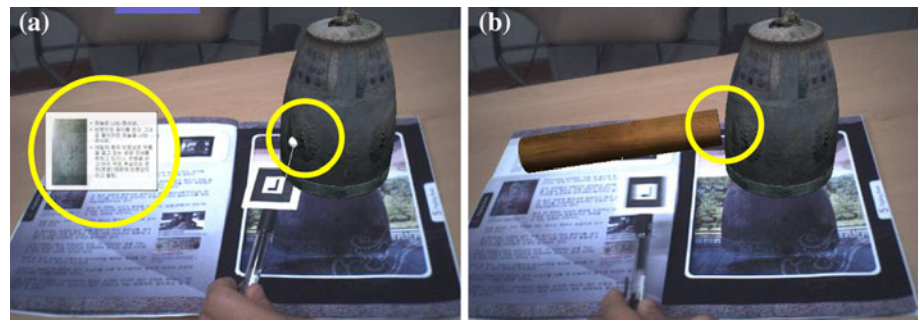

Fig. 14 Virtual button interaction: **a** virtual buttons are projected as a 2D image, which sets the region of interests (*red boxes*); **b** the corresponding chromatic image
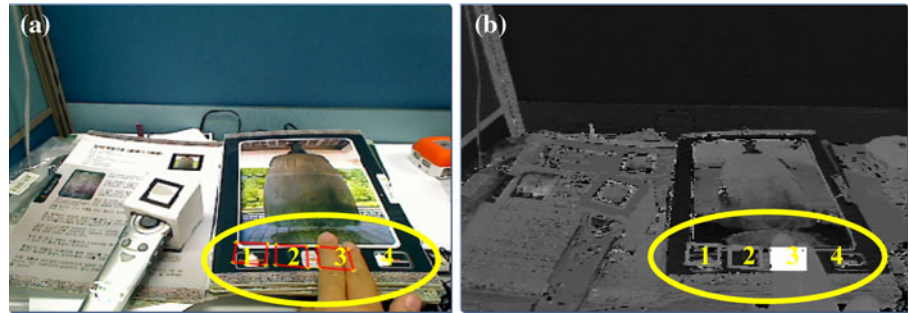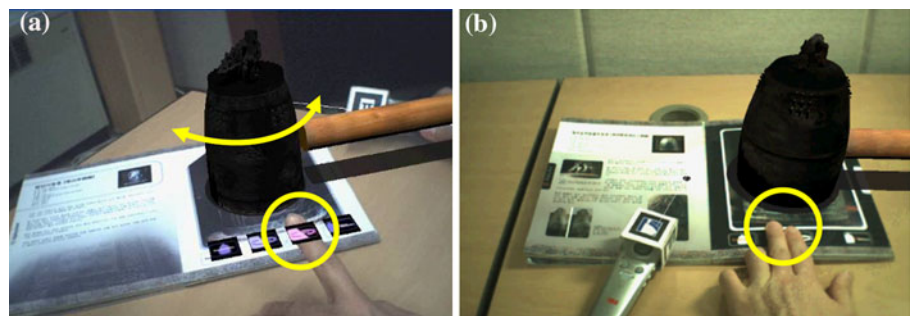

Fig. 15 Selection of virtual buttons, using a user's finger: **a** The second and third buttons are push-button types that trigger the rotation of the 3D model (visible 3D buttons are on the book for confirmation); **b** The fourth button is an auto-releasable toggle button that triggers the hitting bell event (invisible 3D buttons for enhanced naturalism)

5 Discussion and conclusion

5.1 Learning from informal user observation and interviews

This project has collected valuable comments on the Temple Bell Experience Book during many demonstrations and exhibitions over the last 2 years. The section summarizes the learning gained from observation to interviews with informal users.

5.1.1 Input device

This paper discusses a fiducial marker-based manipulation tool as a prototype. Recognition and tracking of the manipulation tool could show better performance at a specific angle between the marker and camera, but this would require user learning time. In addition, prolonged use of the manipulation tool can cause wrist fatigue. As current technology develops, a novel and more comfortable pen-shaped manipulation tool will be available. Additionally, the manipulation tool should be inexpensive so that non-professional users can use it in a home environment, not just as a magnetic tracker-embedded pen in a laboratory.

Although this paper introduces the advantages of direct manipulation with the user's bare hand and with the manipulation tool, another type of user input method is needed to consider the user's requirements. For some readers, a 3D user interface with a manipulation tool would be difficult to control, and a 2D user interface with a mouse input device would be faster and easier to control. Therefore, user preferences and task characteristics for interaction scenarios should be reflected by AR input methods.

5.1.2 Content and interaction

A virtual 3D object should be naturally augmented on figures or layouts of a 2D page. The Digilog Book provides visual special effects and the translation, rotation, and interpolation of a 3D model from a 2D paper with

corresponding position and scale in order to provide more natural augmentations. In addition, the background figure in a book may recede to a vanishing point; in such cases, an augmented 3D object should have a corresponding perspective so that augmentation appears natural.

The Digilog Book should support interesting functions that a paper book cannot provide. The main advantage of a Digilog Book is its provision of visual, audio, and haptic information to a reader. In addition to the methods proposed in this paper, more interesting interactions will be developed, such as explanations of magnified figures or a slice-side view of specific parts of the temple bells.

As a new interface and manipulation method, the Digilog Book can help to pique readers' interest; conversely, however, a user could focus only on the 3D object manipulation rather than on the textual book content. To enforce the Digilog Book's usefulness, manipulation and book content should be arranged in a complementary fashion and offered to the user.

5.1.3 Running environment

The environmental factors of the locations in which the Digilog Book will be used should be considered. Recognition and tracking of fiducial patterns mainly depends on lighting conditions, and also require fine-tuning of the position, distance, and angle between the camera and the fiducial patterns for good performance. One solution is to fix a camera and a book (with printed fiducial patterns) by providing some indications or marks for stable working. For example, in the AR game Eye of Judgment, the camera position is marked as a rectangle on the game pad.

5.1.4 Tracking pattern

Fiducial markers have been involved in many cases of unstable tracking, which occurs when a maker is occluded by a user's hands or a manipulation tool on the book as the user interacts with the augmented objects. To address this issue, many related works have adopted natural feature tracking (NFT). However, depending on the book page's layout and contents, it may be difficult to extract good features and proper tracking for similar patterns (e.g., text majority), dark figures, and reflective material. These are challenging research issues.

5.2 Future direction of the Digilog book

5.2.1 User and environment context awareness

User context awareness, achieved through reader-sensing technologies, could provide personalized interactive stories according to the reader's profile (age, gender, etc.), gesture

recognition, eye tracking, physiological sensing, and more. For example, if a reader's emotional states while reading a book could be recognized, then the book could unfold as a personalized interactive story. The Digilog Book is not fixed like paper books; rather, its story can change to remain interesting and relevant.

While a user reads a Digilog Book, input devices or the user's hands could cast shadows on the book. This reduces naturalism in some situations because it can create conflict with the virtual shadows. To solve this problem, a light-source estimation technique may be useful. This technique measures a light's position in order to generate adaptive shading that mimics the natural shadows in the room where a user is reading the book.

5.2.2 Mobile platform

The Digilog Book viewer is in its early stages of development, and it is currently configured for a desktop computer. As technologies advance in the near future, a portable device for mobile AR will become a general configuration. Using a portable device with any available camera and accessible wireless Internet, a reader could acquire multimedia content from a paper book at any time and in any location. New interaction techniques are also needed to relate to mobile devices with limited display sizes and slightly different input methods.

5.2.3 AR authoring tool

Production of a Digilog Book only can be accomplished by developers and researchers who have a technical background in AR programing. Typical users who lack technical knowledge about programing cannot easily make the Digilog Book. Existing AR authoring tools have been developed to create standard augmented reality applications. The programing-based AR authoring interface primarily uses script languages like programing API and XML for programmers; these are difficult to learn, and they require significant financial and time investment for normal users. Additionally, most users do not specialize in authoring functions for AR-based books. Therefore, this project is pursuing an effective method for producing the Digilog Book that is accessible even to users whose programing knowledge is nonprofessional.

5.3 Conclusion

A Digilog Book stimulates a user's visual, sound, and tactile sensations through the fusion of multimedia content. In particular, the Digilog Book can provide information that conventional paper books cannot. This project has

added the following characteristics to the Digilog Book: multisensory AR content descriptions for updatable AR content; enhanced multisensory feedback to support an immersive experience; a natural, computerized, vision-based bare-hand input method for an interactive experience; and an interaction design. In order to verify the usability of these developments, this project observed and interviewed casual users. The paper also summarized the future directions of the Digilog Book.

As an example of a Digilog Book, this paper introduced the concept and interactive implementation of a temple bell experience book. The book allows users to explore Asian cultural heritage, and it is valuable and applicable because it provides digitalized information about cultural assets. The temple bell experience book gives its readers an interesting and interactive experience with multisensory feedback. Digilog Books are expected to be used in posters, pictures, newspapers, and sign boards.

References

ARToolKit, http://www.hitl.washington.edu/artoolkit. Access date: June 2010

Azuma R (1997) A survey of augmented reality presence. 6, pp 355–385

Billinghurst M, Kato H, Poupyrev I (2001) The magic book: a transitional ar interface. IEEE Comput Graph 25(5):745–753

Chen CH, Hsu MF, Lee IJ, Sun KW, Lin YC, Lee WH (2009) Applying augmented reality to visualize the history of traditional architecture in Taiwan. 22nd CIPA Symposium

Cho K, Lee J, Soh J, Lee J, Yang H (2007) A realistic e-learning system based on mixed reality. In: The 13th Intl conference on virtual systems and multimedia. pp 57–64

Cultural Heritage Administration of Korea, http://english.cha.go.kr. Access date: June 2010

Dünser A, Hornecker E (2007) An observational study of children interacting with an augmented story book. In: The 2nd international conference of e-learning and games (Edutainment). pp 305–315

Grasset R, Duenser A, Billinghurst M (2008) Edutainment with a mixed reality book: a visually augmented illustrative childrens' book. In: International conference on advances in computer entertainment technology (ACE)

Ha T, Woo W (2010) An empirical evaluation of virtual hand techniques for 3d object manipulation in a tangible augmented reality environment. IEEE 3D User Interfaces, pp 91–98

Ha T, Kim Y, Ryu J, Woo W (2006) Enhancing Immersiveness in AR-based Product Design. LNCS (ICAT), pp 207–216

Höllerer T, Feiner S, Pavlik J (1999) Situated documentaries: embedding multimedia presentations in the real world. ISWC (Third Int. Symp. on Wearable Computers), pp 79–86

Irawati S, Ahn S, Kim J, Ko H (2008) VARU Framework: enabling rapid prototyping of VR, AR and ubiquitous applications. IEEE Virtual Reality Conference. pp 201–208

Korean bell's home page, http://library.thinkquest.org/29141//main.htm, Access date: June 2010

Lee G, Kim G (2009) Immersive authoring of tangible augmented reality content: a user study. J Vis Languages Comput 20:61–79

Lee G, Billinghurst M, Kim J (2004) Occlusion based interaction methods for tangible augmented reality environments. In: ACM SIGGRAPH international conference on virtual-reality continuum and its applications in industry (VRCAI). pp 419–426

Lee J, Ha T, Ryu J, Woo W (2009) Development of pen-type haptic user interface and haptic effect design for Digilog book authoring. KHCI, pp 402–405

McKenzie J, Darnell D (2004) The eyeMagic book. A report into augmented reality storytelling in the context of a children's workshop 2003. New Zealand Centre for Children's Literature and Christchurch College of Education

Open Dynamics Engine, http://www.ode.org, Access date: June 2010

osgART, http://www.artoolworks.com/community/osgart, Access date: June 2010

Rauterberg M (2006) From personal to cultural computing: how to assess a cultural experience. In: Kempter G, von Hellberg P (eds) uDayIV–Information nutzbar machen. Pabst Science Publisher, Lengerich, pp 13–21

Saso T, Iguchi K, Inakage M (2003) Little red: storytelling in mixed reality. SIGGRAPH Sketches and Applications, pp 1–1

Scherrer C, Pilet J, Fua P, Lepetit V (2008) Haunted book. ISMAR, pp 163–164

Sellen A, Harper R (2003) The Myth of the paperless office. MIT Press, Cambridge

Shelton B, Hedley N (2002) Using augmented reality for teaching Earth-Sun relationships to undergraduate geography students. First IEEE Intl. Augmented Reality Toolkit workshop. pp 1–8

Shibata F, Yoshida Y, Furuno K, Sakai T, Kiguchi K, Kimura A, Tamura H (2004) Vivid encyclopedia: MR pictorial book of insects. In: the 9th VR society of Japan annual conference. pp 611–612

Stricker D, Daehne P, Seibert F (2001) Design and development issues for archeoguide: a augmented reality based cultural heritage on-site guide. In: International conference on augmented, virtual environments and three-dimensional imaging

Taketa N, Hayash K, Kato H, Nishida S (2007) Virtual pop-up book based on augmented reality. HCI, pp 475–484

The Eye of Judgement, http://us.playstation.com/games-and-media/games/the-eye-of-judgment-ps3.html, 2007. Access date: June 2010

Tosa N, Matsuoka S, Ellis B, Ueda H, Nakatsu R (2005) Cultural computing with context-aware application: ZENetic computer. Lect Notes Comput Sci 3711:13–23

Wagner D, Barakonyi I (2003) Augmented reality kanji learning. In: Proceedings of the 2nd IEEE/ACM symposium on mixed and augmented reality (ISMAR). pp 335–336

White M, Liarokapis F, Darcy J, Mourkoussis N, Petridis P, Lister PF (2003) Augmented reality for museum artefact visualization. In: 4th irish workshop on computer graphics, Eurographics Ireland Chapter. pp 75–80

Wireless presenter: 3M WP-8000, http://www.3m.com. Access date: June 2010

Woods E, Billinghurst M, Looser J, Aldridge G, Brown D, Garrie B, Nelles C (2004) Augmenting the science centre and museum experience. In: 2nd international conference on computer graphics and interactive techniques in Australasia and SouthEast Asia (Graphite). ACM Press, New York, 230–236

Yang H, Cho K, Soh J, Jung J, Lee J (2008) Hybrid visual tracking for augmented books. In: the 7th international conference on entertainment computing. pp 161–166

Zhou Z, Cheok AD, Tedjokusumo J, Omer GS (2008) wIzQubesTM-a novel tangible interface for interactive storytelling in mixed reality. Int J Virtual Real 7(4):9–15