2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)

Visual Augmentation of Printed Materials with Intelligent See-through Glass Displays: A Prototype based on Smartphone and Pepper's Ghost

Frode Eika Sandnes and Evelyn Eika
Department of Computer Science, faculty of Technology, Art and Design
Oslo Metropolitan University
Oslo, Norway
frodes@oslomet.no, evelyn.eika@oslomet.no

***Abstract*—Augmented reality technologies have been applied in educational contexts to enhance traditional textbooks with so-called mixed reality books where static printed context is augmented with dynamic content. These techniques sometimes exploit mobile devices held in mid-air or displays placed behind the textbook. This paper presents a design case outlining a concept for augmenting printed material which operated in the 2D plane by superimposing images on top of the printed material. The contents of a smartphone display are reflected via the printed surface. Several use cases are discussed. The method holds potential for both education and accessibility.**

Keywords-augmented reality, artificial intelligence, sketch, education, learning, accessibility

I. INTRODUCTION

Augmented reality has been applied to many domains including education. One application of augmented reality within education is to enhance the static printed textbooks with dynamic contents making learning more interactive and fun. Although many interactive learning materials exist purely in the digital domain, printed textbooks persist due to their many beneficial characteristics [1] such as portability, availability, and simple navigability.

Typically, an augmented textbook provides 3D enhancements, for example, smoke in a 3D landscape [2, 3]. Digitally augmented books are sometimes referred to as mixed reality books. Some of the literature on mixed reality books report special purpose topic domains such as learning geometric shapes [4] or learning numbers by interacting with 3D characters [5]. This paper presents a design case with several use cases.

Augmentation is sometimes achieved by the means of a smartphone held in mid-air [2, 5]. Currently, all smartphones have cameras, large displays, networking, and processing capabilities making them technically suitable for augmented reality applications; they are also quite affordable and easily available. The augmented and interactive information is thus superimposed on the image captured by the camera. This approach requires complex 3D operations to align the virtual model with the real world captured by the camera. A simpler approach is to place a display behind the book. This display will then show the scene and the augmented information [6]. Although this approach is simpler as the vantage point of the scene remains static, it also means that the user's visual attention may be divided between the real book and the display behind the book.

Most of the literature on augmented reality in education focuses on visual augmentation, but some researchers also have addressed audio augmentation, for instance, allowing readers to touch items in a book that results in audio being played. This has been implemented using RFID-tags embedded in the pages of the book [7]. An overview of various techniques and technologies for augmented reality techniques in education can be found in [8]. Other aspects of augmented reality in education are discussed in [9, 10].

The main contribution of this paper is a concept for realizing visual augmentation of printed material. Instead of augmenting the real world in 3D space, the real world is augmented in 2D on the surface of the printed material. The technique is demonstrated by the means of a concept sketch [11] which is often used to ideate user interface designs [12] and has also been used to ideate immersive 3D-scenes [13]. Several use-cases are explored herein.

II. DESIGN VISION – SEE-THROUGH GLASS

Our vision for the interactive mixed reality textbooks concept is illustrated via the sketch shown in Fig. 1. Printed material is augmented by the means of a fully transparent glass plate-like device. This device can detect the content with which it is overlaying and display information on the glass surface, thereby augmenting information through interactive visual overlays. Clearly, this design vision is not realizable with current technology as it is hard to make fully transparent electronics [14] capable of providing the needed functionality.

III. AUGMENTING PRINTED MATERIALS

As it is not yet possible to accomplish the vision of augmentation though intelligent glass with transparent electronics, three alternative approaches for achieving similar effects are discussed. The first approach is simply implementing augmented reality by the means of a handheld smartphone as illustrated in Fig. 2. Several augmented reality systems are based on this approach (see for instance [2, 5]). The user simply holds the smartphone in the hand while pointing the camera at the book. The content of the book is captured by the camera and displayed on the screen with the augmented information superimposed on the real-world image.

978-1-5386-9269-1/18/$31.00 ©2018 IEEE
DOI 10.1109/AIVR.2018.00063

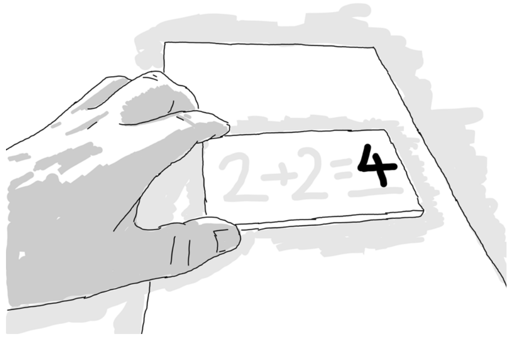

Figure 1. Design vision sketch of augmented glass.

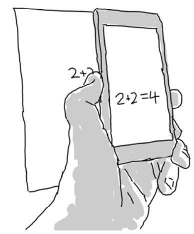

Figure 2. Augmented reality model 1 with handheld smartphone.

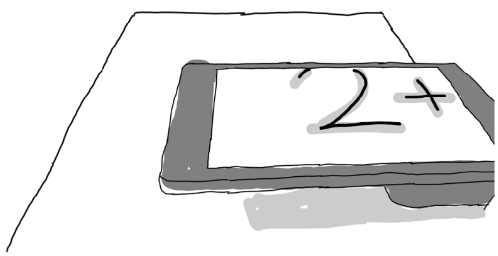

Figure 3. Augmented reality model 2 with smartphone on fixture.

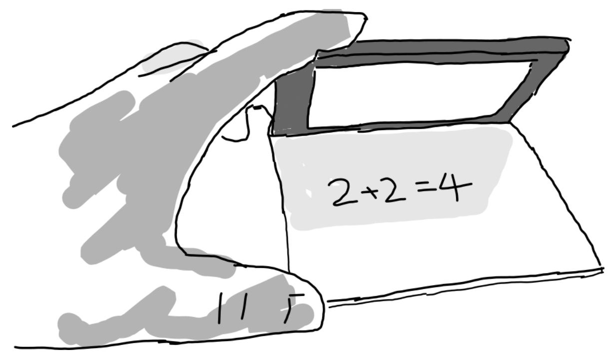

Figure 4. Augmented reality model 3 with smartphone, glass plate, and Pepper's ghost.

There are three main drawbacks of this approach. First, one of the user's hands will be occupied holding the smartphone. This is impractical in many situations such as studying where one hand may be occupied holding a pen while writing and the other hand busy navigating the pages of the book. That is, just like most real-world tasks, studying is often a bimanual activity involving both hands in coordination. Human computer interface designs, however, are often based on one-handed operations. One notable exception is touch typing.

Second, a handheld smartphone will be subjected to shaking. Although several image stabilization algorithms have been proposed [15], computational effort is needed, further complicating the task of recognizing what the camera is pointing at. Third, a handheld phone is positioned in 3D space at some unknown position in an unknown orientation. Image translation and rotation corrections will need to be provided before the content recognition can proceed. This type of 3D augmentation mode was thus deemed too complicated for the target application domain.

The second approach involves a fixture for the smartphone allowing it to rest at a certain distance above the surface of the printed material as depicted in Fig. 3. The smartphone display faces upwards showing the content of the printed material captured by camera pointing downwards. The distance between the printed material surface is sufficiently large for the smartphone camera to capture a sufficiently large image. This approach overcomes three main shortcomings of the handheld version and is used by some commercial digital magnifying glasses (e.g., Eschenbach's SmartLux Digital). Yet, the printed surface is viewed indirectly via the smartphone display and is subject to noticeable lag in the camera display loop. Moreover, the smartphone may cast shadows on the printed material making it harder to read. Although a smartphone usually comes equipped with a flash, these may be too strong for close-up viewing on glossy paper due to potential reflections and specular highlights.

The third augmentation approach attempts to augment the printed material surface directly by projecting the desired image onto the surface. Although this could be achieved with a micro-projector [16], a goal was to achieve the prototype with commonly available components. Micro-projectors have become quite small, powerful, and relatively inexpensive. However, they are still relatively uncommon and not sufficiently small for the given problem. Instead, we designed a prototype using a smartphone exploiting the Pepper's ghost effect [17, 18] (see also [19, 20, 21]). The content on the display is projected onto the printed material surface by a small glass plate which serves as a transparent reflector. Hence, the smartphone display is facing down at a 45-degree angle in relation to the glass plate (see Fig. 3).

By ensuring approximately a 45-degree angle between the observer and the glass plate, and between the glass plate and the display, the contents of the display appear superimposed on the content beneath the glass plate.

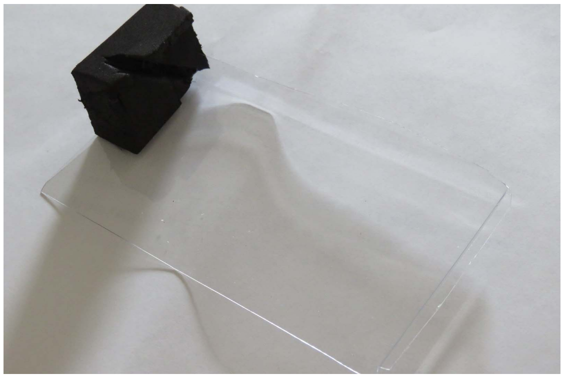

Figure 5. The prototype without the smartphone.

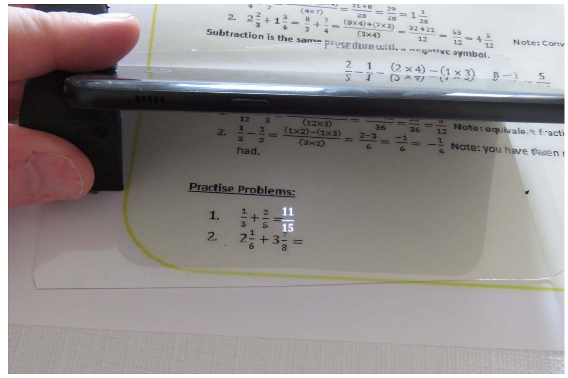

Figure 6. Doing mathematics: summing fractions.

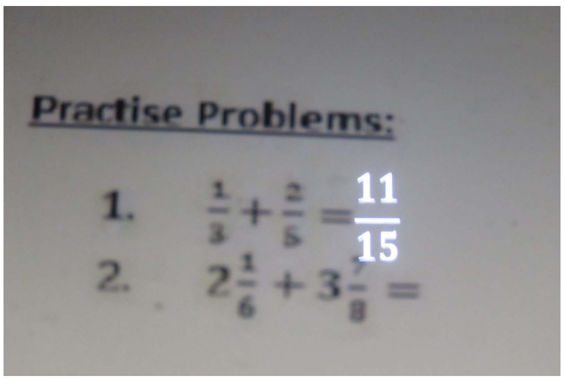

Figure 7. Doing mathematics: summing fractions close up.

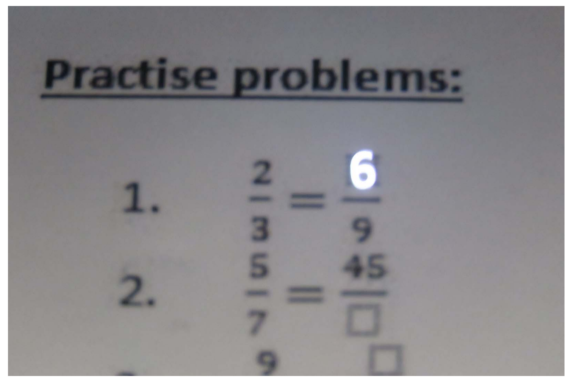

Figure 8. Doing mathematics: converting fractions.

The projection may not be geometrically perfect, but the small size of the display means the geometric distortions have limited effects in practice.

Despite the fact that the smartphone may cast some shadow over the printed surface, it is less problematic than positioning the smartphone directly overhead as with the second approach discussed earlier. If the light source is directly above the printed book, the slanted phone will allow more light to pass through than if the phone is directly overhead. Moreover, the observer will look down on the device at an angle of approximately 45 degrees. The user will thus only see the minimal part of the smartphone since it is viewed from the side along the width, making the smartphone less visually intrusive (see Fig. 6).

Compared to the first two approaches outlined, this design is based on using the camera on the display side of the smartphone instead of the one on the backside. Often the secondary camera on the smartphone has lower specifications than the back camera, yet the cameras on most current smartphone models should suffice as the target is relatively near. The device should be positioned such that the camera is as far from the printed surface as possible. This is especially important if the camera is positioned in the left or the right top corner instead of in the center.

IV. Generating Augmented Visualizations

The information displayed on the smartphone display is viewable by the users via the glass-plate reflection. However, the reflection means that the landscape image is mirrored vertically. Any information to be displayed needs to be mirrored vertically to neutralize the mirroring effect before it can be displayed. Moreover, black does not result in a reflected image, while white provides the most visible reflections. Therefore, the visualizations should have a black background and use white or other colors for clear augmentations. In other words, text should be visualized negatively or inverted.

The visualization approach is similar to what has been used for implementing smartphone-based augmented reality head-mounted displays [22]. This approach reported to be sensitive to the amount of background light where visualizations are less visible in bright environments such as outdoors. However, the applications outlined herein are intended for indoor use under controlled lighting conditions with both lower light intensities and constant light. Educational study activities are usually conducted indoors.

V. Interaction Styles

The user would interact with the device by moving the device around on top of the printed surface. The camera would capture the changes that are passed to the pattern matching engine. This engine would determine what the device is currently overlaying. To help the user control the device, an intelligent viewfinder may be provided. Initially the viewfinder could display a cross that could change into a rectangular box once the cross overlaps an object such as a word or an image.

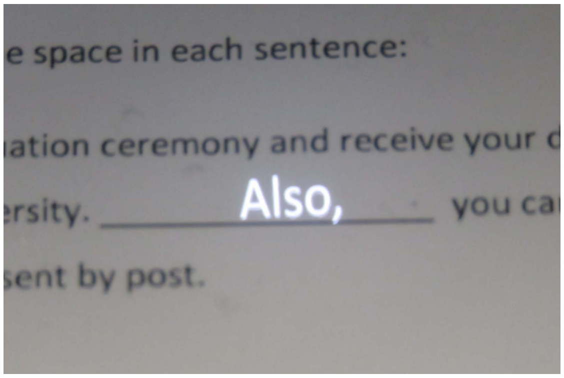

Figure 9. Foreign language learning. Filling in missing words.

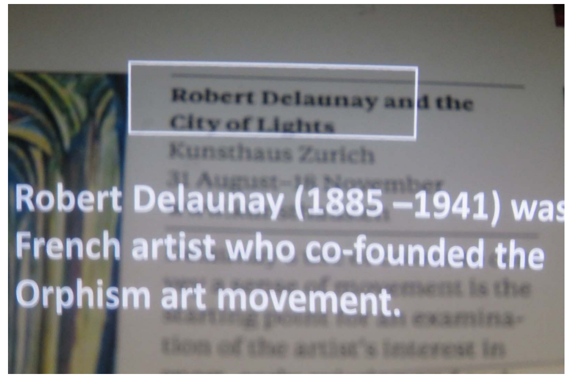

Figure 10. Reference tool for magazine readers. Information about a painter.

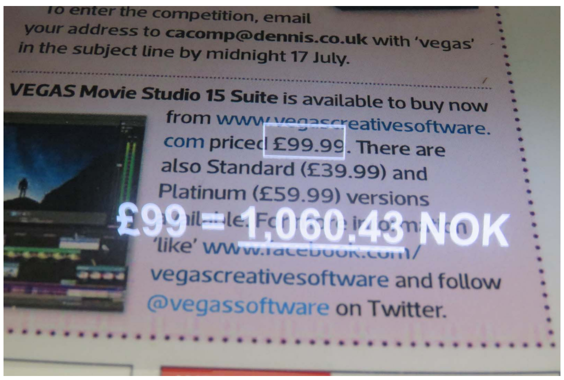

Figure 11. Currency conversion between British points and Norwegian crowns.

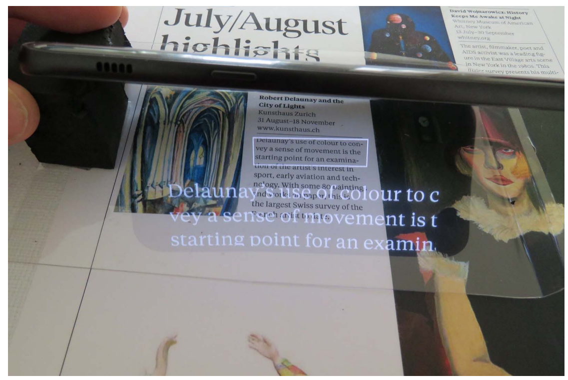

Figure 12. Digital magnifying glass with inverted text.

One could also envisage a setup which does not require the printed material to be known a priori, but rather attempt to analyze what is observed and produce the most appropriate response using sophisticated pattern matching and artificial intelligence, for example, identifying and solving mathematical problems [23] or translating foreign words based on the identified context [24].

VI. Content Analysis and Pattern Matching

Pattern matching is needed to make the device intelligent. The first task is to determine what the device is overlaying. This could perhaps be achieved by analyzing the content of the image captured by the front camera of the smartphone. This image first must be corrected for the geometric distortions as the camera vector is at a 45-degree angle with the printed surface. This could possibly be achieved using basic 3D image transformations [25, 26].

Assuming corrected camera images, there are two ways to perform basic pattern recognition, namely (low-level) image-based pattern matching or (high-level) content-based pattern matching. With image-based pattern matching, the raw image may be compared to a set of existing images of the printed materials using a robust image registration algorithm [27]. This approach is particularly useful if the printed material contains complex layout, images, equations, etc. Alternatively, if the content is mostly text based, the camera image can be converted to text using one of the many state-of-the art optical character recognition algorithms [28] and later robust text-based pattern matching algorithms can be used to detect what the device is overlaying [29]. These approaches assume that the content of the printed material is known a-priori. This is feasible in many situations such as education where a user may "pair" the device with a physical textbook. Note that the current concept prototype does not implement any of the intelligent pattern recognition algorithms described.

VII. AR-Device Concept Prototype

The prototype of the augmented reality device was built using commonly available packaging material, namely, a sheet of clear but stiff plastic with a rounded edge for extra stiffness and for positioning the smartphone and small block of hard polystyrene foam. The plastic sheet was cut into a rectangle with the same dimensions as the smartphone with the rounded edge on the long-side. The block of polystyrene was cut into a cube and a 45-degree slit was cut through one side of the cube for attaching the mobile phone. The cube was then attached to the corner of the plastic plate using double-sided tape as illustrated in Fig. 5. A Samsung Galaxy A5 was used in the prototype. Fig. 6 illustrates how the smartphone is attached to the glass plate.

VIII. Use Cases

This section outlines several use cases for the proposed system. It is envisaged that the system is particularly applicable for education and learning scenarios such as mathematics. Imagine the user is a pupil learning basic mathematics solving textbook assignments. By overlaying the device on top of the mathematical questions in the textbook, the pupil can be given hints for how to solve the problem, corrections, and solutions. For instance, Fig. 6 illustrates how the device is used on top of a sheet with a mathematical tutorial on fractions. The sheet contains the problem 1/3 + 2/5 with a blank space for the student to fill in the answer. By using the device, the user can get a hint of what the correct answer is, namely 11/15. Moreover, the device could also display static or animated depictions of the steps needed to arrive at the answer, which is first creating a common denominator by multiplication and then adding the two nominators. Fig. 7 shows the augmented information close-up. Fig. 8 shows how the device is applied to a different mathematical problem involving fractions, that is, if the fraction is 2/3, what will it be if the denominator is 9. The correct answer is displayed in the blank fill-in box.

Next, imagine a pupil is learning a foreign language, for example, English. Using the device, the user can get immediate help with translation, pronunciation, and explanation of new vocabulary. Fig. 9 illustrates the technique applied to a tutorial sheet addressing English grammar. The pupil is asked to insert the correct linking marker in the bland space, i.e., "You can attend a graduation ceremony and receive your degree certificate from the University. _____ you can graduate in absentia and get your certificate sent by post". The missing sentence opener "Also" is augmented on top of the sheet of paper.

The device can also be used as a general reference tool. Imagine a newspaper or magazine reader who wishes to look up a word, name, or face in a picture. This is illustrated in Fig. 10 where a reader highlights the name of the painter Robert Delaunay in an article published in the international art magazine Apollo, and a brief introduction to Robert Delaunay is displayed on top of the text.

The tool could also be used as a currency converter (see Fig. 11) where the user overlays the device over a price displayed in a foreign currency having the price converted to the home currency with the most updated currency rates. The example in Fig. 11 shows a price in British pounds (£99) in an advert published in the British computer magazine PCPro which is converted to Norwegian crowns (1060 NOK).

The device may also be used as a digital magnifier. The advantage of this approach is that the user will see both the original content and the enlarged content simultaneously eliminating some of the problems associated with confused navigation while panning in zoomed mode. Moreover, as the device can rest on the surface, users with certain type of reduced mobility such as tremor may not experience shaking images. Fig. 12 illustrates digital magnification. A short snippet of text within the augmented rectangle is magnified and displayed on the same page. If the device is moved, the magnified area pans accordingly. In this example the captured image area is inverted so that the text is transmitted as light rather than the background.

Figure 13. Interacting with the device through panning and the use of an intelligent viewfinder.

One could also envisage applications involving printed maps where real-time information could be superimposed on printed maps. An example involving augmentation of maps is displayed in Fig. 13. In this example the user is looking for the main campus of OsloMet on a printed map of Oslo. By moving the device around, the location of the campus is highlighted as a bright disc. By panning the device leftwards on the map, the name of the location is also displayed accordingly on the right side of the disc.

The final use case is the incorporation and augmentation of dynamic social elements into static printed materials. Envisage the user is reading a printed newspaper and may see immediate social media responses in real-time to the printed material as the reader reads the text.

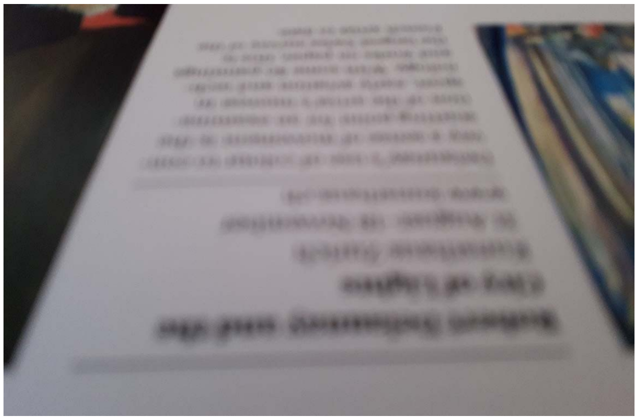

Figure 14. Camera view of the printed surface.

IX. EXPERIENCES

Experimentation with the prototype revealed several challenges. First, the position of the superimposed information obviously depends on the position of the viewers' head. Small head movements can cause the augmented content to easily be misaligned with respect to the printed material. This misalignment effect is larger for content closer to the observer and smaller further away from the observer because of the longer distance between the smartphone display and the point of reflection. The proposed approach is only approximate and not suitable for high-precision tasks. For high precision tasks, a projector may prove useful [30, 31].

Another practical issue is the quality of the images captured by the smartphone camera. The question is if these images are of enough quality to be used by a pattern matching engine. Fig. 14 shows a picture captured using the smartphone camera while positioned in the prototype fixture pointing at the article in the Apollo magazine. There are several issues. First, the image is blurred as it points at a page at a 45-degree angle. Hence, the distances between the camera and the various lines of the page vary greatly. Second, the image is perspective-distorted. Third, the image only captures the page at the horizontal position of the camera in relation to the entire device. Some parts of the page are thus in the camera blind-spot, including content overlaid by the device.

The perspective distortions can be corrected relatively easily using simple image transformations. Next, image registration can be performed using down-sampled and transformation corrected images where the blurring is less of a problem. Down-sampled image registration will also be computationally less intensive than performing the same operation with high-resolution images.

Then, by obtaining a fix on the location and orientation of the camera on the page, it will be possible to know also the content of the blind-spot provided one operates with a content database. Experimentation with real implementations will reveal the actual feasibility of the approach with current smartphone technology.

X. CONCLUSIONS

A prototype of an augmented reality device that superimposes dynamic content onto printed material was presented. Several use cases are discussed in the domains of education and learning, reference tools, and accessibility. Several technical challenges remain unresolved such as the automatic identification of the printed contents. Future work includes user testing to explore to what degree the proposed approach is beneficial to users.

REFERENCES

[1] C. C. Marshall, "Reading and interactivity in the digital library: Creating an experience that transcends paper," Proc. CLIR/Kanazawa institute of technology roundtable, vol. 5, no. 4, 2005, pp. 1-20.

[2] R. Grasset, D. Andreas, and M. Billinghurst, "Edutainment with a mixed reality book: a visually augmented illustrative childrens' book," Proc. International conference on advances in computer entertainment technology, ACM, 2008.

[3] R. Grasset, A. Dunser, and M. Billinghurst, "The design of a mixed-reality book: Is it still a real book?," Proc. 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE Computer Society, 2008.

[4] T. G. Kirner, F. M. V. Reis, and C. Kirner, "Development of an interactive book with augmented reality for teaching and learning geometric shapes," Proc. 7th Iberian Conference on Information Systems and Technologies (CISTI), IEEE, 2012.

[5] A. B. Tomi, and D. R. A. Rambli, "An interactive mobile augmented reality magical playbook: Learning number with the thirsty crow," Procedia computer science, vol. 25, 2013, pp. 123-130.

[6] T. Ha, Y. Lee, and W. Woo, "Digilog book for temple bell tolling experience based on interactive augmented reality," Virtual Reality, vol. 15, no. 4, 2011, pp. 295-309.

[7] M. Back, J. Cohen, R. Gold, S. Harrison, and S. Minneman, "Listen reader: an electronically augmented paper-based book," Proc. SIGCHI conference on Human factors in computing systems, ACM, 2001, pp. 23-29.

[8] M. Kesim, and Y. Ozarslan, "Augmented reality in education: current technologies and the potential for education," Procedia-Social and Behavioral Sciences, vol. 47, 2012, pp. 297-302.

[9] A. Park, K. J. Lee, and F. Casalegno, "The three dimensions of book evolution in ubiquitous computing age: Digitalization, augmentation, and hypermediation," Proc. IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, IEEE, 2010, pp. 374-378.

[10] S. C.-Y. Yuen, G. Yaoyuneyong, and E. Johnson, "Augmented reality and education: Applications and potentials," Proc. Reshaping Learning, Springer, 2013, pp. 385-414.

[11] B. Buxton, "Sketching user experiences: getting the design right and the right design," Morgan Kaufmann, 2010.

[12] Sandnes, Frode Eika, and Hua-Li Jian. "Sketching with Chinese calligraphy," Interactions, vol. 19, no. 6, 2012, pp. 62-66.

[13] F. E. Sandnes, "PanoramaGrid: a graph paper tracing framework for sketching 360-degree immersed experiences," Proc. International Working Conference on Advanced Visual Interfaces, ACM, 2016, pp. 342-343.

[14] J. F. Wager, "Transparent electronics," Science, vol. 300, no. 5623, 2003, pp. 1245-1246.

[15] S. Erturk, "Digital image stabilization with sub-image phase correlation based global motion estimation," IEEE Transactions on Consumer Electronics, vol. 49, no. 4, 2003, pp. 1320-1325.

[16] D. Raboud, T. Barras, F. L. Conte, L. Fabre, L. Kilcher, F. Kechana, N. Abelé, and M. Kayal, "MEMS based color-VGA micro-projector system," Procedia Engineering, vol. 5, 2010, pp. 260-263.

[17] R. Sidharta, A. Hiyama, T. Tanikawa, and M. Hirose, “The Development of Multi-Depth Pepper's Ghost Display for Mixed Reality System,” Proc. 16th International Conference on Artificial Reality and Telexistence--Workshops, IEEE, 2006, pp. 115-118.

[18] D. Sidharta, R., Hiyama, A., Tanikawa, T. and Hirose, M., “Volumetric Display for Augmented Reality,” Proc. 17th International Conference on Artificial Reality and Telexistence, IEEE, 2007, pp. 55-62.

[19] D. Wenig, J. Schöning, A. Olwal, M. Oben, and R. Malaka, “WatchThru: Expanding Smartwatch Displays with Mid-air Visuals and Wrist-worn Augmented Reality,” Proc. CHI Conference on Human Factors in Computing Systems (CHI '17), ACM, 2017, pp. 716-721.

[20] O. Bimber, B. Fröhlich, D. Schmalstieg, and L. M. Encarnação, "The virtual showcase," IEEE Computer Graphics and Applications, vol. 21 , no. 6, 2001, pp. 48-55.

[21] M. Kim, J. Lee, W. Stuerzlinger, and K. Wohn. "HoloStation: augmented visualization and presentation," Proc. SIGGRAPH ASIA 2016 Symposium on Visualization, ACM, 2016.

[22] F. E. Sandnes, and E. Eika, "Head-Mounted Augmented Reality Displays on the Cheap: A DIY Approach to Sketching and Prototyping Low-Vision Assistive Technologies," Proc. International Conference on Universal Access in Human-Computer Interaction, Springer, 2017, pp. 167-186.

[23] B. P. Berman, and R. J. Fateman, “Optical character recognition for typeset mathematics,” Proc. International symposium on Symbolic and algebraic computation, ACM, 1994, pp. 348-353.

[24] J. Yang, J. Gao, Y. Zhang, X. Chen, and A. Waibel, “An automatic sign recognition and translation system,” Proc. Workshop on Perceptive user interfaces, ACM, 2001, pp. 1-8.

[25] R. Raskar, and P. Beardsley, “A self-correcting projector,” Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2001.

[26] T. Okatani, and K. Deguchi, “Autocalibration of a projector-camera system,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 12, 2005, pp. 1845-1855.

[27] L. G. Brown, “A survey of image registration techniques,” ACM Computing Surveys, vol. 24, no. 4, 1992, pp. 325-376.

[28] S. Impedovo, L. Ottaviano, and S. Occhinegro, “Optical character recognition—a survey,” International Journal of Pattern Recognition and Artificial Intelligence, vol. 5, 1991, pp. 1-24.

[29] D. E. Knuth, J. H. Morris, and V. R. Pratt, “Fast pattern matching in strings,” SIAM Journal on Computing, vol. 6, no. 2, 1977, pp. 323-350.

[30] O. Bimber, A. Emmerling, and T. Klemmer, “Embedded entertainment with smart projectors,” Computer, vol. 38, no. 1, 2005, pp. 48-55.

[31] D. Bandyopadhyay, R. Raskar, and H. Fuchs, “Dynamic shader lamps: Painting on movable objects,” Proc. IEEE and ACM International Symposium on Augmented Reality, IEEE, 2001, pp. 207-216.