An Interactive Augmented Reality Coloring Book

Adrian Clark[1], Andreas Dünser[1], Raphaël Grasset[1,2]

[1]The HIT Lab NZ, University of Canterbury, Christchurch, New Zealand

[2]Institute for Computer Graphics and Vision, Graz University of Technology, Austria

ABSTRACT

Creating entertaining and educational books not only requires providing visually stimulating content but also means for students to interact, create, and express themselves. In this paper we present a new type of mixed-reality book experience, which augments an educational coloring book with user-generated three dimensional content. We explore a "pop-up book" metaphor and describe a process by which children's drawing and coloring is used as input to generate and change the appearance of the book content. Our system is based on natural feature tracking and image processing techniques that can be easily exploited for other AR publishing applications.

KEYWORDS: Natural Feature Tracking, Mixed Reality book, interactive AR, edutainment, education, 3D texturing.

INDEX TERMS: H.5.1 [Information Interfaces and Representation]: Artificial, augmented, and virtual realities

1 INTRODUCTION

Since the introduction of the MagicBook by Billinghurst et al. [1], many researchers and companies have created augmented reality (AR) book experiences. Generally such books focus solely on visual augmentation [2,3], whereas some have also explored other features such as audio cues [4] or transitional interfaces [5]. One of the goals of adding such features is to create engaging user experiences, which can enrich the 'reading' made possible by such technology.

However, a very important affordance of augmented reality books is the added interactivity. The most basic interactions are manipulating the physical book as a tangible interface (e.g. rotating or tilting the book pages) to experience the virtual content from different positions or flipping book pages. These techniques are afforded by the physical implementation of an AR book and therefore do not have to be developed separately such as by adding input devices or creating specific setups. Additional tangible interface objects such as cubes can be easily integrated and distributed with an AR book (for example wIzQubes by MXR Corporation[1]).

In this paper we present a new type of interactive AR book experience. We explore how user input through a creative task can bring new interaction opportunities with a book. In particular, we introduce a new technique to engage users in actively creating AR book content in real time using a coloring book metaphor. In our first prototype, users can color the pages of the book, and the system automatically recognizes the pages and the various colors used by the users, and directly maps the colored end results to virtual pop-up scenes and 3D models.

The main contributions of our work are:

- Creating a novel user experience with AR books by allowing readers to interactively create new book content and directly experience their creations
- Detecting and registering a tracking target which may have been modified considerably from the original
- Creating three dimensional scenes and textured 3D models from the users content in real time

In the next section, we present the related work followed by a description of our prototype.

2 RELATED WORK

This work bridges two research fields, education and computer vision in augmented reality. In particular, AR being used as a tool to enable interactive media for the purpose of education. These fields are discussed in the following section.

Multimodality and interactivity has been shown to increase engagement, immersion, and learning. Moreno and Mayer [6] argue that interactivity can promote learning by activating certain cognitive processes. Prior knowledge stored in long-term memory may be activated, organized and integrated with the incoming information. Moreno et al. [7] found that students who are allowed to interact could remember more and transfer what they had learned to new problems better than students who only passively received information. The students work harder to make sense of the material and rate their interest in the material higher.

Dünser and Hornecker [8] concluded that letting children interact with an AR story book may not only be very engaging but could also facilitate recall of story events. In a later study comparing high and low ability readers, Dünser [9] found that low ability readers remembered significantly less information from text-based material, however, there was no difference in recall performance between the two groups with interactive AR content. Therefore AR books that allow readers to interactively engage with content could help students who have problems with the predominantly text-based learning materials used in education today.

There are two methods of computer vision tracking and registration which are applicable to AR books; fiducial registration and natural feature registration. Fiducial registration introduces symbols to a page which have been specially designed to be easily identifiable in vision algorithms [10, 11]. These symbols improve the accuracy and reduce the complexity of registration, but are susceptible to occlusion and modification, and require revision to the original material.

Natural feature registration involves searching for features which are present in an existing image, and describing them in a way which ensures they are identifiable and unique regardless of

---

adrian.clark@hitlabnz.org
andreas.duenser@hitlabnz.org
raphael.grasset@hitlabnz.org

---

[1] http://www.mxrcorp.com/index.php?Itemid=144 retrieved on 26/05/11

changes in viewpoint, orientation, scale, and other transformations [12, 13]. Once these features have been found in both the marker and camera image, the transformation can be calculated [14]. Although natural feature registration is usually slower and may not be as accurate as fiducial registration, there is no need to modify the original material, and due to redundancy afforded by multiple features, occlusion and modification is less detrimental to the success of registration.

3 PROTOTYPE DEVELOPMENT

3.1 Design and Challenge

A natural feature registration approach was chosen to support the augmentation and modification of a real physical book without altering the content of the book. Our first tests with the book demonstrated that greyscale illustrations were providing enough visual features for the tracking.

As our goal is to support modification of the content of the book, the tracking should be sufficiently robust to support changing targets (color changes, added lines, etc). Additionally, as children will not generally follow standard guidelines for coloring a book (only color the predefined areas, don't draw lines over already drawn content, etc.), our tracking should be able to handle errors and uncertainty.

Most natural feature detectors use analysis of the area surrounding each feature point for identification and differentiation of each feature. Even a small change in this area can modify the computed feature descriptor enough to cause matching failure, and with enough failures the calculated transformation will be erroneous, or registration may fail completely. To address this issue, a series of filters were applied to remove as much color as possible from each frame. In the following sections we will describe the different steps of our algorithm.

3.2 Color Removal

First, the sum of the three color components of each pixel was calculated, and the pixels with the maximum and minimum values were found. These pixels represent the lightest and darkest areas in the frame. The maximum of the lightest pixels red, green and blue color components was found, as was the minimum of the darkest pixels.

From these two values, scale and shift values were calculated, which allowed the contrast of the frame to be stretched while maintaining color constancy and ensuring that no color component of the lightest and darkest pixels exceeded the limits of the RGB spectrum.

Once the contrast of the frame was adjusted to maximize the difference between dark and light, the lightest pixel was then used as the white point for color balancing.

With the frame's contrast enhanced and colors balanced, each pixel was evaluated for its grayscale distance. This distance is defined as the distance the pixel is in RGB space from a line running from pure black to a gray value t_1. The value of t_1 determines the "black range", that is the maximum brightness a gray pixel can be before it is no longer considered black.

The grayscale distance of a pixel will be 0 if the pixel falls within the accepted black range; and increases as any color channel within the pixel moves it from the gray line. The distance is normalized on the maximum distance between line and pure red, pure green, and pure blue values.

The final step of color removal involves setting all pixels above a second threshold t_2 to the color white. This allows the user to remove any pixels which are too light to be considered black.

Figure 1 shows the color removal process. Top left shows the original frame, top right shows the frame with a simple grayscale operation, bottom left shows the image after contrast enhancement, color balancing, and gray distance estimation, and bottom right shows the final result after bottom left has been thresholded.
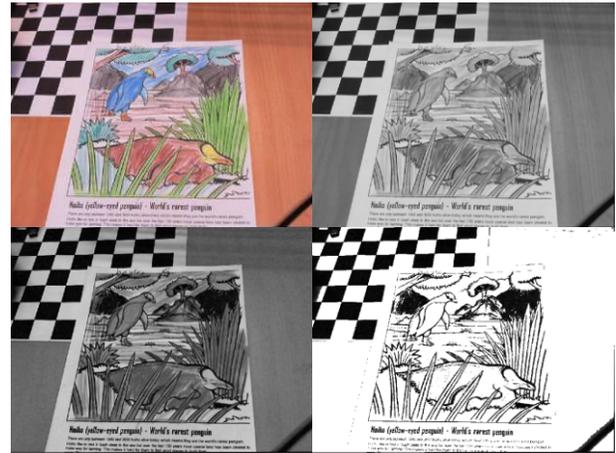

Figure 1. Top left: Original Image. Top right: After standard grayscale operation. Bottom left: After gray distance estimation. Bottom right: Thresholded distance estimation.

3.3 Image Registration

Once the color has been removed from the frame, it can be registered against known pages. The OPIRA registration library [15] is used for registration due to its robustness to perspective distortion and other common image distortions, and the SURF [13] feature detector is used due to its fast computation.

To minimize the effect of registration errors, once a marker has been detected the corner points of the marker are projected into the frame to find their position. If any of the corner points lay outside the frame the application prompts users to ensure the entire page is within frame.

If a marker has been detected but is not entirely within frame, a red rectangle is rendered in the assumed position of the marker. Once the user positions the marker so that it is completely visible, the red rectangle turns blue, and the user is prompted to hold the camera steady for five frames. This delay is to ensure that the marker is stable in the camera view, which reduces the effect of motion blur and reduces the likelihood that an erroneous registration will cause the capture of invalid data.

To further reduce errors, during these five frames Lucas-Kanade Optical-Flow [16] is performed between consecutive frames on a sparse set of features found within the marker. RANSAC analysis [17] is performed across these features, and if the number of corresponding features is below 98%, or if the marker moves out of frame, the five frame counter is reset.

Once five stable frames are recorded, the fifth frame is then rectified to the same orientation and dimensions as the image of the page which was used for registration. The page detection registration process is stopped, and a new registration process for tracking the rectified colored page is started.

3.4 Texture extraction

When a marker has been successfully identified in the frame, the frame can be rectified to recover the new marker. In the rectification the new marker is oriented and resized to match the original, and from there texture extraction is trivial. Any modifications to the original marker can be identified and highlighted, or even applied to three dimensional models with texture co-ordinates which map to the original marker.

4 EDUCATIONAL COLORING BOOK APPLICATION

The New Zealand Department of Conservation provides a coloring book titled "Amazing Animals of New Zealand" to primary school aged children to teach children about animals native to New Zealand. Two pages were chosen for the example application, one featuring a Kiwi, and the other featuring a Penguin.

4.1 Design

To create the virtual pop-up experience for this book, a color mask was created for each page, as shown in Figure 2. Each area designed to pop up as a virtual image layer was colored in a separate color, with the background colored in Magenta.

Figure 2. Left: Original Page. Right: Pop-up Mask

When the application is run, a separate alpha-blended texture is created for each color. The background texture is filled with the original image, such that when the layers "pop-up" it appears the color has lifted off the page. Every other texture is filled with the rectified image captured at the end of the page detection process.

The Y values of each layer when rendered on the marker are determined by the maximum Y value of that color in the mask. As such, masked objects which are lower on the page will be rendered in front of objects which are higher.

An animated model was created for the most important object in the scene, in the two example pages, a Kiwi and a Penguin. These models were texture mapped using the original pages as the reference. When the page is detected, the relevant area of the image is copied into the models texture buffer, such that the model is now textured using the image on the page.
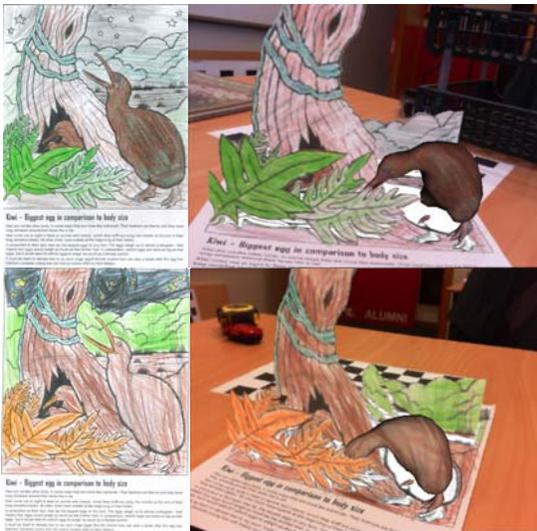

Figure 3. Two examples of the "Kiwi" page. Left: The colored pages, Right: The 3D scenes created in the software

The two pages were given to five children aged between 6 and 8 years old to color in, and their colored pages were used as examples for the application. Examples of these pages are shown in Figure 3 and 4 for the Kiwi and Penguin pages respectively. The colored pages are shown on the left, and the corresponding three-dimensional scenes created by the software are shown on the right.

Figure 4. Two examples of the "Penguin" page. Left: The colored pages, Right: The 3D scenes created in the software

4.2 Performance

To create an engaging AR experience, the application must be able to identify and register pages which have been significantly modified during coloring, capable of running in real time, and able to accurately register the virtual content in the real scene.

From the two pages the children were given to color, five examples were selected for each page. These pages were chosen to cover a wide range of possible cases, including pages where children had colored thoroughly, very lightly, and even colored large areas black. The pages had been colored with pencils, crayon, and even felt tip marker.

With the thresholds t_1 and t_2 set at start-up, 80% of the markers were able to be registered immediately. The remaining 20% required adjustments to the thresholds at run-time, however with these minor adjustments all markers were able to be successfully registered.

On an Intel 2.4Ghz quad core desktop computer capturing images at 640 by 480 resolution, the colour removal and registration steps ran at 25 frames per second, with the majority of the computation required for registration. There was a delay of less than half a second when the new texture was applied to the three dimensional models, but this was due to the time taken to copy the large amount of texture data onto the GPU. After this delay the system ran at 25 frames per second again.

As the rectified marker image is used in AR registration once it has been located, the virtual content was able to be registered to the scene with the same level of accuracy as if the colored images had been captured offline.

4.3 Limitations

The software was designed to be as robust as possible, and easy to use. However due to the nature of the experience, there are limitations.

The software requires the black and white line art as a baseline to identify the original page and calculate the pages orientation. Under different light conditions and for different cameras the

thresholds required to acquire an accurate line art image vary. This could be resolved by an algorithm to automatically adjust the thresholds.

If the pictures are colored with mostly dark or exceptionally saturated color, it is possible that the color removal algorithm will fail completely. In certain examples where large areas of the image were colored in black, color removal failed in these areas and no features were able to be found. The algorithm has been tested and shown to work on images created using color pencils and crayon, however images which have been colored with dark markers or inks will likely fail to register.

4.4 User Feedback

The educational coloring book example was demonstrated at the 2011 International Symposium on Mixed and Augmented Reality (ISMAR'11). Participants were provided with the original black and white colored pages and coloring pencils, and the pages colored by the school aged children were also available for participants who did not want to color themselves.

Over twenty participants chose to color in their own pages, all of which were excited to see their artwork come to life on the page. The users tried different coloring and texturing styles, and even tried to create artistic effects such as having the Swiss flag appear on the penguin's chest, or having the kiwi wear a tuxedo.

Several users deliberately tried to cause the system to fail by adding additional line art details in black, or coloring areas of the image in extremely darkly. Despite this, the system was able to successfully identify and augment every page.

User feedback was overwhelmingly positive. Participants enjoyed being able to see the artwork they had created on the animated three dimensional models. Users thought the simple and natural method of creating artwork was well suited for children, and that it would create an engaging and enjoyable experience for children.

Many users expressed an interest in being able to download the software for themselves for their own children to use.

5 DISCUSSION AND CONCLUSION

In this work we present a new experience utilizing augmented reality enhanced books. Using an educational coloring book example, users are able to color in the pages, and these pages are then recognized by the system and used to produce three dimensional scenes and textured models reflecting the artwork created by the users. This three dimensional virtual content is then overlaid on the real book pages, providing a three dimensional experience using the users own content.

The example applications we created show that the color removal system developed is able to successfully allow automatic registration and texture extraction from coloring book pages, even when the images appear significantly different from the unmodified black and white pages. In the example applications the system was capable of running at real-time frame rates, and with accuracy comparable to capturing the images offline.

At the moment the system requires an initial calibration of two thresholds, t_1 and t_2 to successfully remove the color, however by providing feedback from the registration algorithm these thresholds could automatically be adjusted, removing the need for any manual calibration.

In future we plan to perform a formal evaluation to gauge how these AR books affect engagement in children.

In this initial investigation, as a case study we applied the technology to an educational coloring book. In future we will investigate other applications where the technology may be useful. We believe that this technology would be a valuable tool for artists who wish to create 3D content but lack the necessary computer aided design skills. When combined with an automatic model generation technique [18], there are a number of possible fields which could benefit from this technology for example in architectural design, and rapid prototyping of 3D designs.

REFERENCES

[1] Billinghurst, M.; Kato, H.; Poupyrev, I. "The MagicBook - moving seamlessly between reality and virtuality*," Computer Graphics and Applications*, IEEE, vol.21, no.3, pages 6- 8, May/Jun 2001

[2] C. Scherrer, J. Pilet, P. Fua, and V. Lepetit, "The haunted book", *Proceedings of ISMAR*, 2008, pp.163-164.

[3] T. I. Saso, K. Iguchi, and M. Inakage. Little red: storytelling in mixed reality. *Proceedings of SIGGRAPH '03 Sketches & Applications,* pages 1–1, New York, NY, USA, 2003. ACM.

[4] T. Ha, Y. Lee, and W. Woo. Digilog book for temple bell tolling experience based on interactive augmented reality. *Virtual Reality*, 1—15, 2010.

[5] R. Grasset, A. Dünser, H. Seichter, and M. Billinghurst. The Mixed Reality Book: A New Multimedia Reading Experience. *Proceedings of CHI'2007 Interactivity*, San Jose, USA, April 2008.

[6] R. Moreno, and R.E. Mayer. Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology,* 97(1), 117-128, 2005.

[7] R. Moreno, R.E. Mayer, H. Spires, and J. Lester. The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19, 177-213, 2001.

[8] A. Dünser, and E. Hornecker, An observational study of children interacting with an augmented story book. *Proceedings of Edutainment 2007*, Springer, July 11-13, CUHK, Hong Kong

[9] A. Dünser. Supporting Low Ability Readers With Interactive Augmented Reality. *Annual Review of CyberTherapy and Telemedicine: Changing the Face of Healthcare.* San Diego: Interactive Media Institute, 6, 41-48. 2008.

[10] Owen, C.B.; Fan Xiao; Middlin, P. "What is the best fiducial?" *The First IEEE International Augmented Reality Toolkit Workshop*, pages 98-105, 2002.

[11] Fiala, M. "Comparing ARTag and ARToolkit Plus fiducial marker systems," *IEEE International Workshop on Haptic Audio Visual Environments and their Applications*, pages 147-152, 1-2 Oct. 2005

[12] Lowe, D.: 1999, "Object recognition from local scale-invariant features", *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, Vol. 2, pages 1150-1157

[13] Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding (CVIU),* Vol. 110, No. 3, pages 346-359, 2008

[14] Zhang, Z.: 2000, "A flexible new technique for camera calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, pages 1330-1334.

[15] Clark, A., Green, R. and Grant, R.: 2008, "Perspective correction for improved visual registration using natural features", *Image and Vision Computing New Zealand, 2008. IVCNZ 2008*. 23rd International Conference, pages 1-6.

[16] Lucas, B. D. and Kanade, T.: 1981, "An iterative image registration technique with an application to stereo vision", *Proceedings of the 1981 DARPA Image Understanding Workshop*, pages 121-130

[17] Fischler, M. A. and Bolles, R. C.: "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography", *Commun. ACM* 24, volume 24, no 6, pages 381-395.

[18] Clifford, R.; Clark, A.; Rogozin, M.; Billinghurst, M.; "An augmented reality approach to architectural building design". Accepted for publishing in *12th Annual Conference of the New Zealand Chapter of the ACM Special Interest Group on Computer-Human Interaction*, 2011