Augmenting physical books towards education enhancement

G. Margetis[1], A. Ntelidakis[1], X. Zabulis[1], S. Ntoa[1], P. Koutlemanis[1] and C. Stephanidis[1,2]
[1]Foundation for Research and Technology – Hellas, Institute of Computer Science
N. Plastira 100, Vassilika Vouton, GR-700 13 Heraklion, Crete, Greece
[2]University of Crete, Computer Science Departmnent
{gmarget, ntelidak, zabulis, stant, koutle, cs}@ics.forth.gr

Abstract

This paper presents an interactive desk that augments physical books that are placed upon its surface with multimedia content and interactive applications. Such content is dynamically displayed in augmentation to the currently open page of the book, that is, aligned in real-time with its 2D orientation upon the desk. The rendered applications are controlled by the users with the use of a stylus, both through contact with the book or desk, as well as, through a small vocabulary of gestures performed with the stylus. The evaluation of the accuracy, robustness, and performance of the proposed computer vision modules supporting this interaction are reported through quantitative experiments. In addition, the system usability was validated and the suitability of educational applications was explored through pilot applications, which include music and digital animation on pages, content-based multimedia presentation, context-based online search, as well as, note-taking through handwriting.

1. Introduction

The idea of digitally augmenting physical paper has stimulated scientific interest since the early 90's. DigitalDesk [14] and its successor EnhancedDesk [10] were two of the pioneering digital paper augmentation applications that ignited inspiration for a number of different approaches and extensions in the next two decades. The basic idea of these systems was the use of Computer Vision technologies for physical objects recognition and localization, enabling intuitive interfaces that seamlessly linking physical and digital documents.

In the education domain, augmented reality comprises a contemporary medium providing novel environments for students to learn abstract concepts. To this end, an avenue of research has focused on physical book augmentation. Examples of physical book augmentation, [7, 2, 9], aim to enhance the conventional reading experience with 3D graphics and digital content.

In this work, a smart desk that augments physical books placed upon its surface with multimedia and interactive applications, related to the content of the open book page, is proposed. Such content is dynamically displayed in enrichment to the currently open page of the book, and is aligned, in real-time, with its 2D orientation upon the desk. In addition, 3D gestures are supported enabling the users to intuitively interact with the physical books, the rendered digital content, and applications. Currently, interaction is supported through the use of an infrared *lightpen*: a stylus shaped, LED based, infrared light (IR) emitter. A prototype of the system has been developed, to evaluate the proposed system and supporting modules. This prototype showcases a book about the work of the Nobel laureate poet Odysseus Elytis. Accompanying applications focus on the presentation of the poet's work and target the effortless and entertaining education of users.

The remainder of this paper is organized as follows. In Section 2, related work is reviewed. In Section 3, the modules of the proposed system and the architecture that interrelates them is presented. In Section 4, the findings of the accuracy, performance, and usability evaluation of the proposed system are reported. Section 5 concludes this paper and provides directions for future work.

2. Related work

Systems employing physical book augmentation emerged a decade ago. In [1], a system combining an interactive multi-track audio environment with text and images printed on traditional paper pages was presented. Each page of the book was augmented with ambient sounds, which gently adapted to the moods and scenes, each time the user touched pertinent regions of the page.

More recent approaches combine immersive environments with augmented reality. MagicBook [2], augments physical books with 3D graphics and moving avatars through handheld augmented reality displays (HHD), giving to the reader the sense of living pages. In Live Book [9], a mixed reality book service that provides physical book augmentation using a small projector is presented. Live Book supports a more natural interaction with the system through the users' fingers, which are recognized using an IR camera and active illumination, but interaction is limited only inside the open page of the augmented book. PapierCraft [11] is a system using pen gestures on paper to support active reading. It allows users to carry out a number of actions, such as to copy and paste information from one document to another.

In this work, an augmented reality system providing content-sensitive paper book augmentation is proposed that combines book and page recognition as well as lightpen position identification. The proposed system displays digital information, either upon or laterally to it, related to the page of a physical book that is currently open. At the same time, users also interact with the system through the use of a lightpen. The displayed applications are controlled by the users with the use of a small set of gestures, carried out with the pen, which provide helpful facilities for the reader.

3. System description

The system employs a desk above which a projector and two calibrated camera pairs, one conventional and one sensitive only to IR light, have been mounted. The first is used for book page recognition and localization; the two

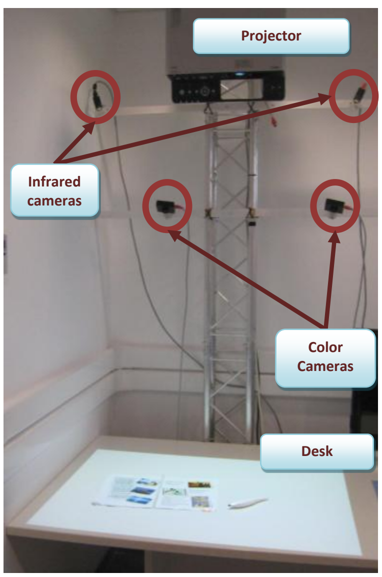

Figure 1. System setup: shown are a desk, upon which books are placed, two pairs of cameras overlooking the desk (one color and one sensitive only to IR), and a wall-mounted projector above the desk which creates the augmented display.

(or even more) cameras are combined to fully cover the workspace in high resolution. The second is used in the stereo localization of the light pen. The projection covers the desk's surface (see Figure 1).

To support content-sensitive augmented interaction with a physical book a module recognizes the book and the page that it is open at and, also, estimates the location and orientation of the book on the desk. A second module supports user lightpen based interaction by providing events that (a) detect and localize its contact to the desk or predefined page regions and (b) recognize context-specific gestures upon the desk, book, and space above the desk.

The system architecture is illustrated in Figure 2. On the left, the "Virtual camera" module combines images from the color camera pair in a single image that covers the workspace in sufficient resolution for robust book page recognition. This image is utilized by the "Book page recognition and pose estimation" module, to provide the id, location, and orientation of open book pages. On the right, images from the IR stereo pair are utilized to track

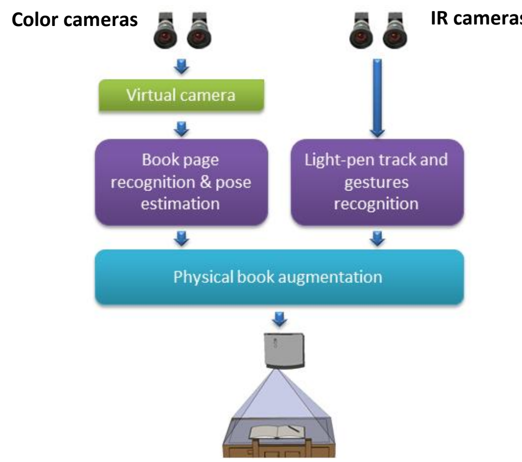

Figure 2. System architecture.

the infrared lightpen device in 3D, detect its contact with book regions or desk locations, as well as, recognize gestures performed with the lightpen. The output of the modules constitutes the basic input for the context-aware "Physical book augmentation" module, which displays visual content superimposed and juxtaposed to the physical book pages. In the following, a detailed description of these components is provided.

3.1. Book page recognition and pose estimation

The purpose of this process is twofold. First, it provides context information to the system regarding the book page which the user is engaged at and, thereby, the book content that is pertinent at the each moment in time. Second, it provides geometric information about the 2D pose of the book upon the desk, that is, its 2D location and 2D orientation which, in turn, facilitates book augmentation (see Section 3.5). To ensure robust page recognition and accurate book localization, a composite view of the desk is initially created, which emulates a virtual camera of high resolution and a wide FOV. This composite image is, then, provided as input to the book recognition and localization process.

3.1.1 *Virtual Camera*

Image based recognition and pose estimation benefit from sufficient resolution, in terms of robustness and precision respectively, which are typically hindered by the wide breadth of the workspace. We, thus, use an additional camera to meet both requirements. The “Virtual camera” module composes input from the two (or more if available) cameras into a panoramic image that covers the entire workspace in sufficiently high resolution. The resultant image is orthocanonical and frontoparallel to the desk surface, to facilitate coordination of physical workspace, book, and projection coordinates. The recognition system treats the input as a single, conventional image.

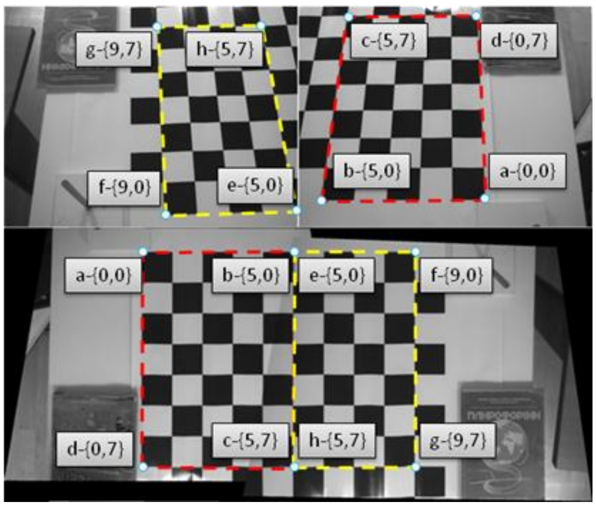

Figure 3. Virtual camera calibration and operation. Left: Regions of an actual calibration image pair with checkerboard portions selected. The pixel coordinates of four points in each image {a-d} and {e-h} are corresponded to the local coordinate system of the checkerboard to determine a homography for each input image. Right: using these homograhies, input images are warped into the composite image.

Figure 4. Image obtained during system operation showing that the composite image correctly emulates a camera for surfaces on the desk (but not for the hand which lies above it).

Images acquired from the cameras are warped and placed into the panoramic image using homography matrices [8]. For each camera image, we calculate a homography using four pairs of 2D points between the desk plane (physical coordinates) and the image coordinate plane (pixel coordinates). A checkerboard aligned with the desk is used as a calibration target and constitutes the physical reference to the desk's coordinate plane. To prevent aliasing, the sampling resolution of the homography is set to be dense enough so that every input image pixel is sampled at least once. Given checker size and the homography upon the desk plane, coordinates on the composite image can be scaled to metric. Note that the composite image will respect the geometry of the physical world only for surfaces on the desk plane. However, this is sufficient, as the system’s operation requires that the book is placed upon it. For this reason, during calibration, the checkerboard is elevated with respect to the desk surface to a height approximately equal to the height of a book.

The, offline, calibration process with the pair establishment of 2D points of the physical plane and the image plane is illustrated in Figure 3. In Figure 4 the composite image obtained during system operation illustrates that the correctly emulation of a camera only for surfaces on the desk.

3.1.2 *Book page recognition, 2D localization, and orientation estimation*

The book recognition module detects the presence of known 2D books on the desktop, recognizes them, and estimates their location and orientation on the desktop coordinate frame. This is achieved by using the recognition system in [12]. This system additionally provides the location and orientation of the book. Using these location and orientation estimates, the boundaries of the book on the desk are predicted. As the composite image is orthocanonical with respect to the coordinate frame of the desk, these predictions correspond to metric coordinates, thus availing an estimation of the physical spatial extent of the book on the desk. Finally, locations and orientations are tracked through Kalman filtering [3], [6] to cope with transient recognition failures.

3.2. Lightpen 3D interaction device

The proposed system employs a calibrated stereo pair that is sensitive only to IR light for the estimation of the 3D location of the infrared LEDs of pens. Visible light is filtered out by two high-pass 850nm filters mounted in front of each camera lens and, thus, only the infrared component of scene illumination is imaged by the cameras. Ideally, only the lightpen's LED should appear in the images, but in practice noise may also occur by other IR light sources, such as sunlight entering from the room's windows and corresponding reflections. The workspace is 3D and is limited to the volume above the interactive desk, which is covered by the field of view (FOV) of both cameras. For convenience in gesture recognition (Section

3.3), the 3D coordinates are provided in the reference frame of the desk, where z=0 is its surface and x and y axes aligned with its edges.

The proposed module utilizes the images acquired by the stereo pair to estimate the 3D locations of the pen's LED in real-time by employing, in each acquired image pair, the following operations: (1) Detection of LED in both images of the stereo pair; (2) Estimation of the 3D LED location; (3) Tracking of the 3D LED location.

Typically, in step 1, LED appears as a bright dot against a dark background in images. The dot is extracted as blobs using image thresholding and a Connected Component Labeling process [4]. Due to sunlight noise, spurious blobs may occur in the images. To deal with these blobs, masks are applied on the images of the stereo pair that restrict the search to the workspace above the interactive desk. Furthermore, the appearance of the LED's blob has been a priori modeled as to its mean intensity, compactness, as well as minimum and maximum areas in the images. Based on this modeling, spurious blobs are filtered out and not detected as LEDs. Upon detection, dots are represented by their centroids. If a LED is detected in both images, then, in step 2, the two dots corresponding to the lightpen are corresponded. The locations of their centroids are inputted to a triangulation procedure [8], along with the camera calibration data, to estimate the 3D location of the LED center. Finally, in step 3, a 3D tracker is incorporated in the stereo system. The tracker uses 3D Kalman filtering [3], [6] to better estimate and also compensate for transient disappearances of dots, i.e., due to noise or occlusions. The latter is achieved by inputting to the tracker, upon disappearance of a 3D point, the estimates of its next predicted state for a small sequence of frames (~10).

Note that the two cameras are calibrated using conventional calibration techniques and a checkerboard as the calibration target. However, in order for this target to be visible by all cameras, the scene is illuminated with infrared light during calibration.

3.3. 3D gesture recognition

The 3D output of the previous module is used to capture trajectories of the lightpen during the system's operation. A distinct trajectory begins upon the first detection of a lightpen in the images. The trajectory ends when there is no 3D estimate, meaning that the lightpen was switched-off or exited the workspace. Immediately afterwards the 3D points in the trajectory are gathered and the acquired 3D trajectory constitutes a gesture to be recognized. The supported 3D gestures are left click, right click, line and circle. Gesture recognition is resolved using rules that take into account the duration of the gesture, whether the pen maintains a stable locus within the duration of the gesture, and whether the path of the gesture better fits a line or a circle in the 3D space. Based on the 3D information, the pertinent UI events that are generated are attributed regarding their occurrence upon the desk, the book surface, or the volume above the desk.

The *left* and *right click* gestures are defined only upon the desk surface and, thus their detection requires that its trajectory points occur (approximately) upon it. A *click* gesture is highly localized and its detection assumes that the variance of the trajectory locations is smaller than threshold T. A *left click* is differentiated from a *right* one with respect to the duration of the gesture, with *left* having a relatively longer duration.

For the *line* gesture the Euclidean distance L_l is calculated between the first and the last position of the trajectory points and compared to the length of the actual trajectory. If the two lengths are approximately equal a *line* gesture is detected. The *line* gesture can be used either on the desk surface or in the entire volume of the workspace.

For the *circle* gesture a plane is fitted among the trajectory points, using the RANSAC framework [5]. Their orthogonal projections on that plane are calculated and they are rotated to be parallel to x-y plane. The z-coordinate is eliminated, and the problem is reduced to fitting an ellipse in 2D. A *circle* gesture is recognized if the fitted circle exhibits small deviations from a perfect circle. As in the case of the *line*, the events inform as to if the gesture occurred upon the desk or above it.

3.4. Projector-camera calibration

Projection is modeled through a homography that predicts the coordinates, in the projected display, that will illuminate a given point on the desk. The homography is calculated a-priori, using point correspondences that are established from the following procedure: the projector projects points on the desk and the user is instructed to switch on the lightpen at their locations. Given the high quality optics of projectors lens, distortion of the projector is assumed to be negligible. More than four pairs (~10) are used to achieve an accurate homography estimation. In this way, given a locus on the desk, visual content can be projected upon it and implement the augmentation precisely. This allows an estimation of 3D location of the book on the desk surface, projecting them to pixel coordinates of the desired IR camera for which we estimated a homography and transfer them to pixel coordinates of the projected image.

3.5. Workspace and physical book augmentation

This component is responsible for providing context-aware interactive content to the user. Technically, augmentation is supported by the projector-camera calibration. Given the coordinates of the book or the lightpen in the desk coordinate frame, this calibration is used to predict the coordinates of the projector pixel that will illuminate the corresponding region or point of interest.

The content-sensitive digital content provided can be classified in the following categories (asset types):

- Images, optionally followed by informative text
- Videos that have been stored in the system
- Images from Google and videos from YouTube that are being collected at run-time according to the user's interaction
- Ambient light control

In more detail, every book that is included in the system's recognition library has been stored in digital form (PDF) and it has also been enhanced with hot-spot areas that play a role in the user-page interaction. The classification of book pages is contained in a XML description file, in which the interactive areas (hotspots) and their types are defined. Every hotspot is declared by a number of coordination points, defining the hotspot's bounding shape and asset type that this hotspot denotes.

Each page has a unique id and is accompanied by its digital image of 600dpi, enabling the system to extract any part of the page annotated by the user through a circular gesture. The extracted page part is presented near the book and the user is able to email it or post it to his/her Facebook or Twitter account.

Every asset type is assigned a number of accepted gestures (e.g., point, circle, etc). The gestures supported have been selected to be intuitive, facilitating the learning of system operation and increasing its usability.

When a page hotspot is engaged by the user's interaction, the system evaluates that input and, according to the type of the asset that is represented, it selects the appropriate supportive applications and displays them on or near the active book's page. For example, in Figure 5 left the user has clicked the hot-spot area signifying an Elytis' particular poem and a YouTube video slideshow related to that is displayed near the open page. If the user rotates and moves the book on the desk the displayed application follows it, keeping its initial position and orientation in respect to the book (Figure 5 right).

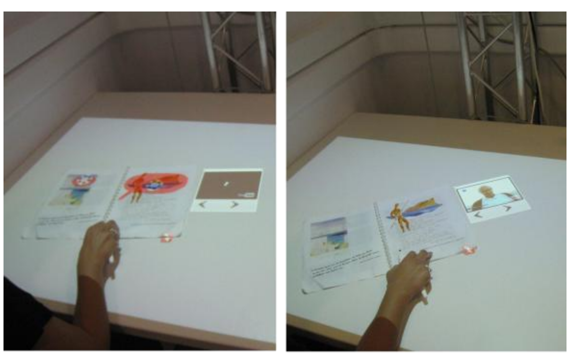

Figure 5. Physical book augmentation. Left: The user clicks an area of the page and triggers a video gallery right laterally to it. Right: While the book is rotated and moved on the desk, the augmentation retains its relative position to the book.

In the previous example, video results can be controlled using a small set of gestures. For example, if a user wants to see the next video, she can virtually draw a horizontal line from left to right on the table, while if she desires to see the previous video, a horizontal line should be drawn to the opposite direction. In contrast, if she wants to clear the digital content that is currently being displayed, she can virtually draw a horizontal line in the air.

In addition to the aforementioned functionality, the presented system provides a note-taking facility, by writing on a free area of the table using his/her infrared lightpen device. The user is also able to associate her notes with parts of the open page by circling them. Furthermore, she is able to choose the ink color and size, while she can email or post the notes to her Facebook and Twitter account.

Additionally, users can move any application that is displayed on the desk's surface and place it wherever convenient using the infrared lightpen.

When the user finishes her interaction with the book she performs a circle gesture in the air notifying the system to turn on the already dimmed lights.

4. Results

In the context of the presented work a performance assessment and usability evaluation were performed. Next sections discuss some important results.

4.1. Performance and accuracy

The system utilizes a PC with an Intel i7 processor, an NVidia Geforce GT330 graphics card and 8 GB of RAM. The stereo pair subsystem comprises of 2 Point Gray Research, Flea2 cameras (1280 by 960 pixels) operating at 30 fps. The lightpen module operates at ~29-30 fps.

For the deployed setup, we evaluated the accuracy of the triangulation by capturing trajectories of straight lines of an illuminated pen on the desk (x-y) plane and above it (x-z) plane. We used a ruler to drag the pen, measure the ground truth line length and compare it with the trajectory length of the triangulation estimate. The mean L2 error of the system was ~2 mm, subject to accurate extraction of the LED centroid on the stereo pair images.

The accuracy of the point transfer between the IR cameras to the projector coordinates was also evaluated. Five users where asked to illuminate the lightpen on 72 points projected by the projector on the desk surface. Pixel coordinates of the extracted centers of the lightpen led extracted from an IR-camera were transferred to the image projected by the projector device using a homography.

Ideally, the estimated transferred points would occur on the points projected by the projector. The mean L2 error of this point transfer was ~2.12 pixels with a standard deviation of ~1.42 pixels. This accuracy proves to be sufficient both in technical and in usability matters.

The Virtual Camera module employs two Point Gray Research DragonFly2 cameras (1024 by 768 pixels) cameras operating at 30 fps. The generation of the composite image is implemented on GPU with performance subject the graphic card's capabilities. In this setup, the resolution of the composite image is restricted to 2048 by 1080 pixels, at a rate of ~20-29 fps.

The system was also evaluated as its recognition rate for illustrated books of more than 300 pages. The recognition of pages was tested initially under no or mild occlusion conditions (~20% of the book pages were occluded, typically by hands holding the book). Under these conditions, no page mismatch was observed, thus providing a 100% confidence regarding the recognized pages. This high confidence is attributed to the uniqueness of the SIFT descriptor and the abundance of features in the image that is dominated by the book.

Recognition was also tested under harder conditions by occluding the book with the user hands. It was found that to retain the same confidence in page recognition the book pages cannot appear occluded by more than ~70% in the image. In such cases, the matching percentage is particularly low and the system notifies the user that the pages are not clearly visible to properly operate. In the given setup, the error in book localization accuracy has a mean average of 5 mm, when under only mild occlusions. This accuracy is attributed again to the abundance of features that are employed in the homography computation. In the experiments, the average retrieval was below .4 sec for the evaluated books.

4.2. Usability evaluation

Following an iterative design approach, the system was developed incrementally, having each version evaluated and the system refined according to the evaluation results. Once a fully-functional interactive prototype of the system was completed, it was evaluated through the heuristic evaluation method [13]. The evaluation yielded seven usability problems, five of which were characterized as relatively minor usability problems and two of which as relatively major.

Table 1. Evaluation participants' characteristics

Gender		Age		Expertise	Techn.	Us/lity
M	40%	*20-30*	26.7%	*Low*	33.3%	33.3%
F	60%	*30-40*	60%	*Medium*	40%	46.7%
		40-50	13.3%	*High*	26.7%	20%

The issues that were identified were addressed and the new version of the system was evaluated through user testing with 15 participants. The user group characteristics, in terms of gender, age, technology expertise and usability expertise are described in Table 1.During the test, the participants were welcomed, and introduced to the system and the evaluation goal. Then, they were asked to carry out specific tasks with the system, including to launch and interact with digital content (music, images, video) related

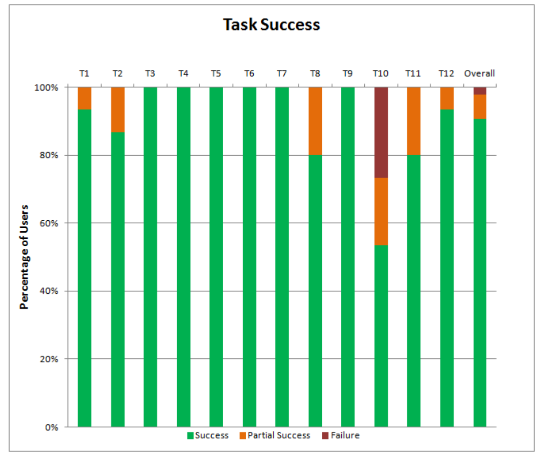

Figure 6. Task success rates.

to the physical book, use the light pen to take notes and share them through e-mail. Task accomplishment was rated as success, partial success (e.g. if the participant asked for help) and failure. Figure 6 presents an overview of the percentage of users who were successful, partially successful or unsuccessful for each task and the average success percentage for all the tasks.

Then, users were asked to complete a subjective evaluation questionnaire, aiming to assess their overall user experience. In more details, the questionnaire involved 12 statements, to which the participants had to specify their agreement on a scale from 1 (strongly disagree) to 5 (strongly agree), as well as two open-ended questions asking participants to identify additional functionality that would be desired and to provide additional comments on the system. Five of the questions were related to the interaction with the system (e.g. if it responded timely, if it was awkward, etc.), while the remaining seven questions aimed to assess the user experience and find out whether users would actually use such a system if it was commercially available and for which purposes (e.g. education, entertainment). In summary, users were satisfied with the system (avg.: 4.26), with the interaction (avg.: 3.97) and with the overall user experience (avg.: 4.47). An initial comparative analysis (see Table 2) among the various user groups indicated that female users as well as younger users tended to be more enthusiastic, while no effect was found by users' technology or usability expertise. A formal comparative analysis per user category, involving chi-square and t-tests will be carried out in future work, through a more extensive evaluation with larger user sample size.

Table 2. Comparative analysis per user category

	All	M	F	20-30	30-40	40-50
Overall	4.27	4.03	4.43	4.52	4.15	4.29
Interaction	3.97	3.73	4.13	4.40	3.78	4.00
UX	4.48	4.24	4.63	4.61	4.41	4.50

Finally, participants were debriefed and a detailed discussion was carried out on the specific tasks that incommoded them as well as additional features that would be desired.

In summary, the evaluation resulted in a list of 25 usability problems ranging from aesthetic to minor (e.g. usage of not intuitive icons, lack of explanatory text along with multimedia content) as well as a list of 10 additional system features that would be desirable. Furthermore, most of the participants stated that they were unfamiliar with the light pen, which they found somewhat cumbersome for notes-keeping. This is also depicted in the satisfaction questionnaire results, where the interaction category scored lower than the user experience category, due to the interaction with the lightpen.

5. Conclusion

This paper has presented an augmented reality system aiming to enhance the process of reading physical books. A fully-functional prototype of the system was tested in terms of performance and usability. The performance test confirmed that the presented system is robust and reliable enough for practical use. The usability evaluation has on the one hand identified some issues that need to be addressed and on the other hand suggested some features that could be included in future versions, such as: have the system read aloud selected passages of text, include educational games, provide users with browser functionality that could be used to retrieve additional information if necessary, support background customization through predefined themes, and support touch-based interaction. Overall the user-based evaluation has yielded positive results regarding the users' satisfaction and their overall user experience. In addition, discussion with participants during the debriefing session, as well as, their answers in the subjective evaluation questionnaire indicated that the system is expected to be very useful in the education domain, allowing students to work in a more effective, efficient and pleasant way. Future work will address the usability issues that were identified, include the additional features that were requested and carry out further usability evaluations with users, focusing on assessing the system with students, using school textbooks augmented with educational-related multimedia and applications (e.g., videos presenting experiments, educational games, etc.).

Acknowledgements. This work is supported by the FORTH-ICS internal RTD Programme 'Ambient Intelligence and Smart Environments'.

References

[1] Back, M., Cohen, J., Gold, R., Harrison, S. Minneman, S.: Listen reader: an electronically augmented paper-based book. In Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '01), pp. 23-29. ACM, New York, NY, USA (2001)

[2] Billinghurst, M., Kato, H., Poupyrev, I.: The MagicBook - moving seamlessly between reality and virtuality. Comp. Graph. App. 21(3), 6- 8 (2001)

[3] Bishop, C. M.: Pattern Recognition and Machine Learning. Information Science and Statistics, 2nd printing. Springer, Singapore (2006)

[4] Chang, F., Chen, C., Lu, C.: A linear-time component labeling algorithm using contour tracing technique. Comput. Vis. Image Und., 93 (2), 206-220, Elsevier, (2004)

[5] Fischler, M., A. and Bolles, R., C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. of the ACM 24 (6): 381–395, (1981)

[6] Ghahramani, Z., Hinton, G.: Parameter estimation for linear dynamical systems. Technical Report University of Toronto, CRG-TR-96-2, (1996)

[7] Grasset, R., Dunser, A., Seichter, H., Billinghurst, M.: The mixed reality book: a new multimedia reading experience. In: CHI '07 Extended Abstracts on Human factors in computing systems (CHI EA '07), pp. 1953-1958. ACM, New York, NY, USA (1958)

[8] Hartley, R., Zisserman, A.: Second Edition Multiple View Geometry in computer vision. Cambridge University Press, Cambridge (2003)

[9] Jeong, H.T., Lee, D.W., Heo, G.S., Lee, C.H.: Live Book: A mixed reality book using a projection system. In: 2012 IEEE International Conference on Consumer Electronics (ICCE), pp.680-681, (2012)

[10] Kobayashi, M., Koike, H.: EnhancedDesk: Integrating paper documents and digital documents. In: 3rd Asia Pacific Computer Human Interaction (APCHI'98), pp. 57–62, IEEE (1998)

[11] Liao, C., Guimbreti, F., Hinckley, K., Hollan, J.: Papiercraft: A gesture-based command system for interactive paper. ACM Trans. Comput.- Hum. Interact., 14 (4), pp. 18:1-18:27 (2008)

[12] Margetis, G., Zabulis, X., Koutlemanis, P., Antona, M., and Stephanidis, C.: Augmented interaction with physical books in an Ambient Intelligence learning environment. Mult. Tool. Appl. Published on-line: 10 January 2012

[13] Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people (CHI '90), Jane Carrasco Chew and John Whiteside (Eds.). ACM, New York, NY, USA, pp. 249-256

[14] Wellner, P.: Interacting with paper on the DigitalDesk. Commun. ACM 36 (7), 87–96 (1993)