Multimed Tools Appl
DOI 10.1007/s11042-011-0976-x

Augmented interaction with physical books in an Ambient Intelligence learning environment

George Margetis · Xenophon Zabulis · Panagiotis Koutlemanis · Margherita Antona · Constantine Stephanidis

Abstract This paper presents an augmented reality environment for students' improved learning, which is based on unobtrusive monitoring of the natural reading and writing process. This environment, named SESIL, is able to perform recognition of book pages and of specific elements of interest within a page, as well as to perceive interaction with actual books and pens/pencils, without requiring any special interaction device. As a result, unobtrusive, context - aware student assistance can be provided. In this way, the learning process can be enhanced during reading with the retrieval and presentation of related material and, during writing, by the provision of assistance to accomplish writing tasks whenever appropriate. The SESIL environment is evaluated in terms of robustness, accuracy and usability.

This paper presents an augmented reality environment for students' improved learning, which is based on unobtrusive monitoring of the natural reading and writing process. This environment, named SESIL, is able to perform recognition of book pages and of specific elements of interest within a page, as well as to perceive interaction with actual books and pens/pencils, without requiring any special interaction device. As a result, unobtrusive, context – aware student assistance can be provided. In this way, the learning process can be enhanced during reading with the retrieval and presentation of related material and, during writing, by the provision of assistance to accomplish writing tasks whenever appropriate. The SESIL environment is evaluated in terms of robustness, accuracy and usability.

G. Margetis (✉) · X. Zabulis · P. Koutlemanis · M. Antona · C. Stephanidis
Foundation for Research and Technology – Hellas (FORTH), Institute of Computer Science,
Heraklion, Crete GR-70013, Greece
e-mail: gmarget@ics.forth.gr

X. Zabulis
e-mail: zabulis@ics.forth.gr

P. Koutlemanis
e-mail: koutle@ics.forth.gr

M. Antona
e-mail: antona@ics.forth.gr

C. Stephanidis
e-mail: cs@ics.forth.gr

C. Stephanidis
Department of Computer Science, University of Crete, Heraklion, Crete, Greece

Published online: 10 January 2012

Keywords Augmented reality · Augmented books · Page recognition · Gesture recognition · Pose estimation · Handwriting · Ambient intelligence

1 Introduction

In recent years, the emergence of Augmented Reality (AR) and Ambient Intelligence (AmI) is driving a transition from traditional human computer interaction to natural and intuitive interaction with everyday things [1]. AmI is often claimed to bring a significant potential in the domain of education [2, 10], while Information and Communication technologies already permeate educational environments in many ways. They can play an important role in education by increasing students' access to information, enriching the learning environment, allowing students' active learning and collaboration and enhancing their motivation to learn [8].

In today's educational settings, reading and interacting with physical book is an important part of the learning process. Currently, the issue of physical vs. electronic books is under debate both in general and in the context of education. The work presented in this paper proposes a combination of both physical and electronic books through augmentation and natural gesture interaction. Research work has been conducted in the past towards enhancing the book reading process through AR. For example, in [17] the authors developed a system which enhances a physical book, thus providing to young readers the opportunity to interact live with its content, by placing fiducial markers on it and seeing augmented 3D graphics, through a handheld display. Reading and interacting with this augmented book provides users with a multi-sensory experience, which is difficult to achieve through other technologies.

However, work conducted so far in applications of AR techniques to physical books have focused on touch interaction, and have not taken into account hand writing. On the other hand, as it is stated in [31] "during the act of writing, there is a strong relation between the cognitive processing and the sensorimotor interaction" (p. 398), enhancing thus the students' understanding and learning ability.

This paper presents a learning support system based on augmenting the physical book and integrating handwriting. As a result, students can be further enhanced with related material when reading, and can be assisted to accomplish their writing tasks whenever appropriate. The underlying approach is based on the recognition of book pages, as well as of specific elements of interest within a page through the intuitive and unobtrusive monitoring system of students' gestures and writing, thus facilitating context – aware content sensitive assistance.

2 Related work

2.1 Ambient Intelligence in learning environments

According to Cook et al. [11], any smart environment can be adequately decomposed in four fundamental layers: physical, communication, information and decision. Each layer performs a different role in the environment, facilitating diverse operations and addressing specific requirements. To this end, some fundamental requirements have to be addressed by a smart environment architecture in order to enable ambient intelligence in a classroom.

An example of Ambient Intelligence approach in the classroom environment is the Smart Classroom [42], which facilitates collaborative learning among college students using pervasive computing technology. In more detail, it integrates handheld devices, such as Personal Digital Assistants (PDAs), with fixed computing infrastructures (e.g., PCs, sensors, etc.) in a wireless situation-aware network in the sense that any device can capture different situations in a classroom dynamically to form ad hoc networks facilitating both student-student and student-instructor collaboration. Features such as the aforementioned are considered to be fundamental for developing intelligent environments that aim to augment the educational process in the classroom.

Another approach towards Ambient Intelligence in the classroom is the ClassMATE smart classroom framework [25], aiming to provide a robust and open ubiquitous computing framework suitable for a school environment that: (i) provides a context aware classroom orchestration based on information coming from the ambient environment, (ii) addresses heterogeneous interoperability of Ambient Intelligence services and devices, (iii), facilitates synchronous and asynchronous communication, (iv) supports user profiling and behavioral patterns discovery and (v) encapsulates content classification and supports content discovery and filtering.

The system presented in this paper interoperates with the ClassMATE smart classroom framework in order to provide context – aware assistance to the students, enhancing thus their learning process during reading with the retrieval and presentation of related material and, during writing, through the provision of assistance to accomplish their writing tasks whenever appropriate.

2.2 Visual 3D pose estimation of a stylus in 3D

The introduction of highly unique and invariant features [28] has boosted the ability to recognize images. The approach proposed here exploits well established planar image recognition methods in the provision of a natural interface through the manipulation of a pen or stylus. This approach falls in the domains of ubiquitous computing and ambient intelligence, where physical objects are blended in interaction comprising “tangible interfaces" [21].

Recent trends in human computer interaction indicate the plausibility of tangible and natural interaction through simple hardware. However, to the best of our knowledge a visually tracked stylus in 3D has not been proposed as a means of interaction. The most relevant work is [19], where a pipette is tracked in from a single view to indicate points on a plane. Markers are utilized to track the wand, at a relatively slow rate (4 Hz).

The need for a pointing device as an interaction item is discussed in [43], where a PDA is visually tracked to emulate a wand in an augmented reality environment. Multiview tracking of markers on the PDA provides pose information for augmenting a virtual wand, whereas the system described in this paper employs a physical wand. This need is also found in efforts to capture human pointing gestures, i.e., [36].

To find its pose (location and orientation), the wand it is reconstructed as a straight 3D line segment, from its simultaneous appearance multiple images. It is avoided to employ conventional stereo [47] for this task, as it typically fails to reconstruct thin objects, due to lack of pixel support. In contrast, images are processed to model the wand as a line segment in the image that represents the projection of the wand in it. Then the problem is modeled as finding the 3D line segment that better explains the obtained observations. Conventional and well-performing approaches in multiview geometry [18] are followed to combine multiple images of a line wand into an estimate of its pose. An early approach to this problem is proposed in [6], which was formulated for 3 views, while the approach proposed here works for any number of available views. In [39] lines are only affinely reconstructed. More pertinent is the approach in [7], which is iterative as the goal there is to match as many segments as possible and thus difficult to parallelize. The approach in [32] assumes the existence of multiple interconnected line segments to detect endpoints, information which is not available in the case of the system described here. The work in [35] detects line segment representations of rooftops from multiple aerial images in essentially the same mathematical manner, however there line segments are terminated based on information from groups of line segments (the edges of rooftops), while in the case of the reported system the endpoints have to be inferred from the observation of a single segment.

2.3 Physical paper augmentation

Since the early years of 90's the idea of digitally augmenting physical paper was intriguing enough to trigger the first research efforts in this direction. For example, DigitalDesk [46] and its successor EnhancedDesk [23] performed physical paper augmentation with technology, offering interaction via touch.

Since then, more sophisticated AR solutions have been proposed by exploiting the means offered by immersive environments and high quality 3D graphics. For example, MagicBook [8] provides augmentation of physical books with 3D graphics and moving avatars through VR glasses, giving to the reader the sense of living pages. The basic interaction technique in such environments is touch. Pointing and writing in augmented reality environments has also been studied, but the majority of research work is based on proprietary technological artifacts like light pens, pen with pads, haptic devices, etc. ([14, 41]).

Finally, the availability of new digital pens capable of capturing marks made on paper documents has led to the development of various systems exploiting user annotations on physical paper in order to assist users' reading process. For example, the Anoto system [3] combines a unique pattern printed on each page with a digital pen to capture strokes made on paper. PapierCraft [26], on the other hand, proposed a system which used pen gestures on paper to support active reading and allow users to carry out a number of actions, such as to copy and paste information from one document to another. Finally, the Paper++ system [29] uses documents overprinted with a non-obtrusive pattern that uniquely encodes the x-y (and page) location on the document; this code can then be interpreted by a specialized pen when it comes in contact with the paper. In such approaches specialized devices (such as the digital pens) are necessary. In order to achieve context-awareness systems such as the Paper++ have to also employ specialized paper besides the digital pens.

To the best of the authors' knowledge e-paper augmentation with plain pens or pencils has not been reported yet, since existing approaches use specialized pens and/or papers. Furthermore, the proposed work introduces context-aware paper augmentation, i.e., handwriting augmentation related to the specific page content, with the use of plain paper. At the same time, existing approaches to context-aware physical paper augmentation had to rely so far on specialized papers with overprinted patterns e.g., [29].

In more detail, this paper introduces an augmented reality environment for students' improved learning, which is based on unobtrusive monitoring of the natural reading and writing process. This environment, named SESIL, does not require any special writing device in order to monitor the student's gestures and handwriting, as it is able to perceive interaction with actual books and pens/pencils.

Furthermore, SESIL provides context – aware help to the students, launching educational applications on a nearby display. SESIL can thus enhance the learning process by unobtrusively and naturally providing additional information related to the current student's activity.

3 System architecture and Implementation

In this section, the architecture of the proposed system is presented, followed by detailed descriptions of the implementation of its comprising software modules.

A functional prototype of SESIL has been deployed on the smart school desk presented in [4], which enhances a standard school desk. The enhancements constitute two additional pieces of hardware (see Fig. 1). First, it is a piece of furniture that frames a 27 inch wide touch screen. Second, a triplet of calibrated cameras that overlook the desk's workspace and supporting computer vision methods, which serve the same purpose. Both enhancements are aimed to support the interaction of the user with the system. These methods are employed by SESIL to recognize the book page that is open on the desktop and also estimate the pose of pen, pencil or stylus that is freely manipulated by the user (henceforth this object is called a stylus).

The smart desk is intended here as an instance of an interactive artifact in the context of an Ambient Intelligence learning environment (e.g., a classroom or the student's room at home). However, the SESIL environment could be easily deployed also in different spatial arrangements, such as at larger interactive surfaces featuring with display and multi-touch capabilities (i.e., for use in a library or at an exposition) provided that the cameras are positioned appropriately.

3.1 Architecture

The presented system interoperates with the ClassMATE smart classroom framework [25]. The ClassMATE's architecture, as presented in Fig. 2, introduces five fundamental components, which are presented in more detail in Table 1.

Figure 3 illustrates the architecture of SESIL, which is designed to facilitate information flow in real-time, so that system interactivity is served. The interactive touch screen operates as an autonomous, conventional component and its design and implementation are reported [33].

The images acquired by the cameras are employed to locate the book on the table, recognize the particular page that the book is open at, estimate the 3D orientation and location of the stylus, and capture contact information between the stylus and the book or the desk. A software module [49] is employed to acquire images from the cameras and place them in a shared memory in the host computer, where SESIL processes can directly access them. As vision processes read these images, they can simultaneously access them in this memory.

The first module in the pipeline contains two vision processes, which read the acquired images. The first module, "book & page recognition" (see Section 3.2), uses only the image from the central camera to identify and locate the page that the book is open at (or the id of

Fig. 1 SESIL deployed on the augmented desk

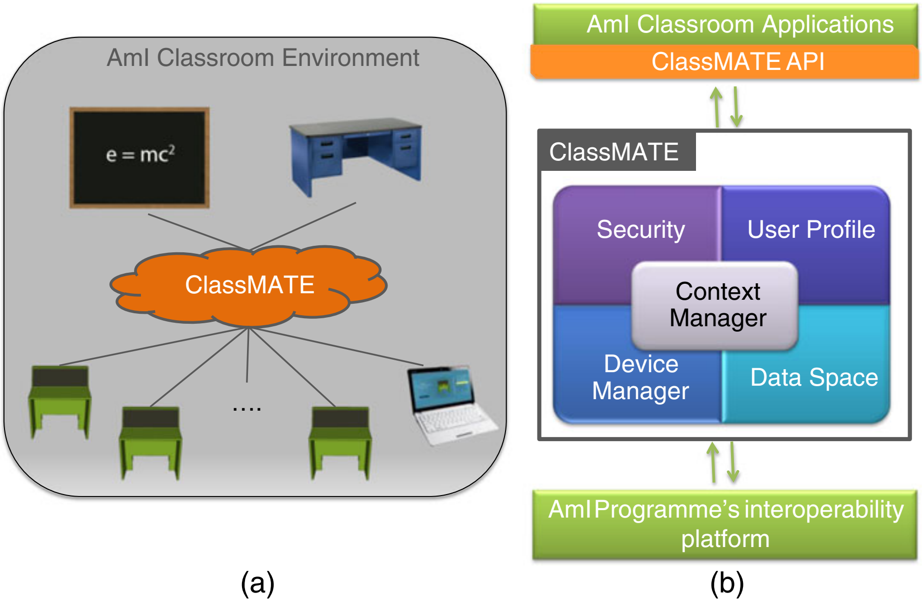


Fig. 2 **a**) ClassMATE provides a complete monitoring and coordination infrastructure for an AmI Classroom Environment. **b**) The ClassMATE architecture

the front/back cover, if that is the case). It outputs the id and the location of the book page(s) in the coordinate frame of the physical desktop. The second, "stylus pose estimation", accesses all three images from the cameras to locate the stylus in 3D at the same coordinate

Table 1 Summary of ClassMATE's architecture components

Context Manager	Platform Expert	Provides access to the wide-variety of platform-specific functions in a platform-independent manner
	ClassMATE Event System	Defines a hierarchy with specialized event types forming the ClassMATE event type system and implements the essential mechanisms for their distribution
	Artifact Director	A context aware module that orchestrates each artifact
	Class Orchestrator	Controls every aspect of the classroom in a top-level
	State Serialization Manager	Manages application's state serialization and deserialization
	Migration Processor	Facilitates the application migration from the current local artifact to a remote node
Device Manager	Multitouch Device Manager	Enables multitouch interaction schemes
	Book Localizer	Determines current context of use (e.g. currently studied course)
User Profile Manager		Collects personal data associated with a specific user (both static and dynamic)
Data Space Manager		Provides an abstraction layer between the applications and the physical storage layer and encapsulates a filtering mechanism for personalized content delivery based on user needs and preferences
Security		a special-purposed module used to handle user or system access-related requests

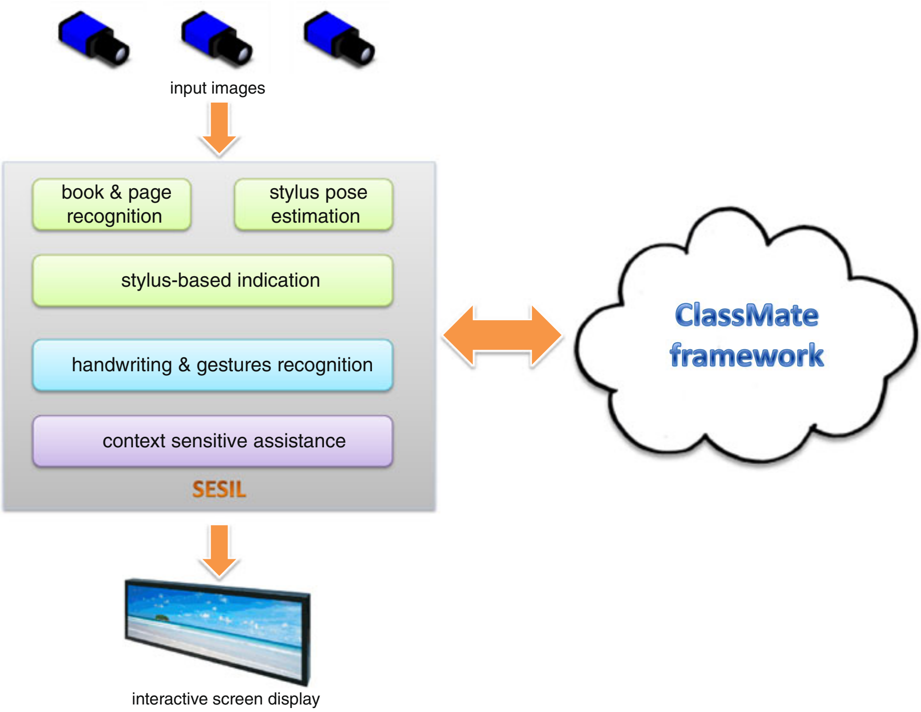


Fig. 3 SESIL's architecture

frame. When the stylus is in contact with either the book or the desk surface, the "stylus-base indication" process communicates an event which contains the coordinates of the contact point. When this contact occurs on the book, these coordinates are also computed in the intrinsic coordinate frame of the book (i.e., the image coordinates of the electronic representation), so the user's pointing by contact action is associated with the electronic representation of the content that this action points at.

The "gestures and handwriting recognition" module takes as input the output of the aforementioned modules and provides gestures and handwritten text recognition. It also propagates the recognized page to the next module.

Finally, the "content sensitive assistance module" elaborates the output of the previous modules, in order to decide the appropriate complementary information needed by the user while reading and writing on the book. This information is displayed on a nearby screen. For testing purposes, SESIL has been integrated in the augmented school desk described in [4], which has been provided with the required front cameras.

3.2 Book and page recognition

A vision process detects the presence of a known book on the desk, localizes it and recognizes the particular page that it is open at, providing a robust context cue to the application. An electronic representation of the appearance of each page is assumed, which is typically acquired from the PDF version of the book (using [5]), or by manually scanning its pages. The outcome of the whole process is demonstrated in Fig. 4.

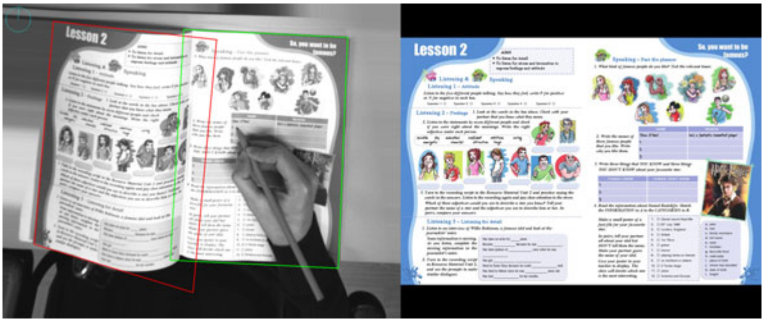


Fig. 4 Recognition of book pages and their localization. On the left, a grayscale version of the image acquired from the central camera is shown. On the right, the matched page in the electronic representation. The left image is superimposed with the estimates of the spatial extent of each page (green & red quadrilaterals) as well as book orientation (cyan circle)

In the electronic representation, each page is represented in a planar, frontoparallel posture that is devoid of any distortions (i.e., projective and lens distortions that are introduced in photographs of the page). A description of each page is a priori obtained by detecting SIFT features [28] in the electronic representation and storing them in a list, in no particular order. At run time, the page at which the book is open at is recognized, as the PDF page that maximizes the matched matching features with the currently acquired image. Thereafter, through these matches, the book is localized using a robust estimation process.

When a frame is acquired, it is initially warped to the plane of the desk, providing image W, within which SIFT features are detected. These features are compared with the ones already in memory, on a per page basis. The page that provides the greatest hit ratio is considered as the matching page and, thereby, matching features establish correspondences between the acquired image and the electronic representation of the particular page. To optimize the search, consecutive and preceding pages to the current are considered first, and if they provide a very high ratio are immediately considered as matching.

The locations of the point correspondences between W and the scanned pages provide a basis for the localization of the book on the desk. For each frame, the homography that warps each imaged page (out of 2) to the scanned one is estimated through a RANSAC procedure [13], along with a subset of confidently correct correspondences (inliers). Using the 3×3 homography matrix, the corners of each scanned page are found in the acquired image. These 4 projections mark the spatial extent of the book in W. Procrustes analysis [9] is then employed, upon the inlying correspondences to estimate it's the 2D orientation of the book within W. Since W is a homography of the desktop, the spatial extent and orientation of the book are directly output in desktop coordinates.

The process is parallelized on the GPU of the computer. At each frame that a book is found, it outputs an event with the ids of the two (or one, i.e., if the cover is matched) imaged pages along with the coordinates of each page in the desk's coordinate system, as well as the orientation of the book (see Fig. 4).

3.3 Stylus pose estimation

A vision process estimates the position and orientation (3D pose) of a colored stylus in the space above the desk, while it is manipulated by the user. To provide the estimation, the process utilizes 3 synchronous views that overlook the space desk, providing images I_i, $i=\{1,2,3\}$, at each "frame". The stylus is reconstructed as a 3D line segment. The outcome of this process is demonstrated in Fig. 5.

Segmentation Each image I_i, is binarized into image M_i to extract the stylus. Segmentation is based on a pixelwise color matching metric [44]. In M_i, a pixel is 1 if the corresponding pixel in I_i matches the stylus color, and 0 otherwise. Binary image M_i typically contains noise mainly in the form of false positive pixels, due to accidental color matching of book regions with the stylus. At the same time, in the particular images acquired by SESIL there exist occlusions of the stylus, at least, due to the fingers manipulating it.

Detection The wand is sought in each M_i, using the Hough Transform (HT) [12], which yields 2D line l_i. Input to HT is provided by a thinning of M_i. which is necessary as the stylus may appear in varying width in the image due to the effect of projective distortion. Due to occlusions and segmentation errors, the wand may not appear as a single segment along l_i, while spurious blobs may also be encountered. A traversal of M_i is performed along l_i and connected components are labeled. Very small segments are attributed to noise and are disregarded. Size-dominant segments along l_i are grouped by adapting the line grouping metric in [27] for the 1D domain of line l_i. Intuitively, the process detects the longest line-shaped object of a particular color present in the image. It is represented as a line segment, s_i, after its endpoints in the image are located. The process is demonstrated in Fig. 6.

Pose estimation The 3D line L where the stylus occurs in 3D space is estimated as the one minimizing reprojection error to the observed line segments in images I_i. Endpoint estimates are then, similarly, obtained from their 2D observations. The input to this process is the lines l_i and segments s_i obtained in the previous step. Outlier elimination upon these inputs is crucial as, due to occlusions and segmentation errors, the object may not be fully visible in all views.

Given lines l_i in each view, the stylus is reconstructed in 3D, as follows (see Fig. 7). For each view, the plane defined by l_i and camera center K_i is considered. Thus per view-pair $j=\{1, 2, 3\}$ a 3D line C_j is obtained by intersecting the two corresponding planes. Lines C_j are combined in the output estimate of the stylus, 3D line L, as in [45]. The endpoints of the 3D segments found in images T_i, are projected and grouped onto E to find the 3D endpoints of the line segment. In Fig. 7, the geometry of the above method is illustrated. At each frame upon which the stylus is found, an event containing the two endpoints e_1, e_2 in the coordinate system of the desk is output.

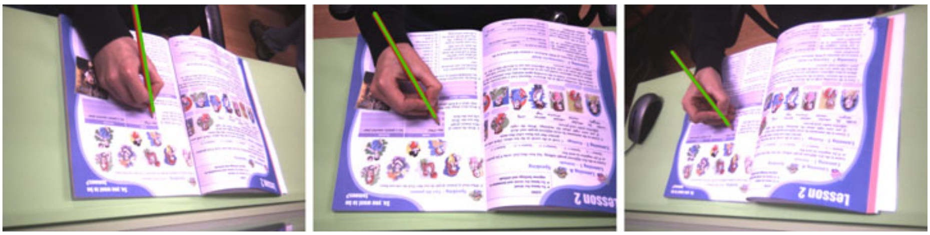


Fig. 5 Stylus pose estimation. Original images acquired from the three cameras that overlook the desk. In each image, the pose estimate (location and orientation) of the stylus is superimposed as a green line segment, by projecting this estimate back to the original image

Fig. 6 Line segment detection. Original image I_i (left), M_i (middle), and M_i with the detected line l_i and the line-segment s_i superimposed in red and green colors, respectively. Notice that the pen segment appearing in front of the user's fingers is grouped with remainder stylus representation as a line segment

Motion estimation is finally applied to improve the accuracy of pose estimation and correct errors, i.e., when the wand is transiently lost or pose estimation is extremely inaccurate. The trajectory of the wand is tracked over time in the 6D space (3D location and orientation) using a Kalman filter [16] which, for each measurement, provides a pose estimate. As this estimate exhibits increased accuracy with respect to the measurement, it is adopted to refine the motion estimation of the stylus. The details of this formulation can be found in [48].

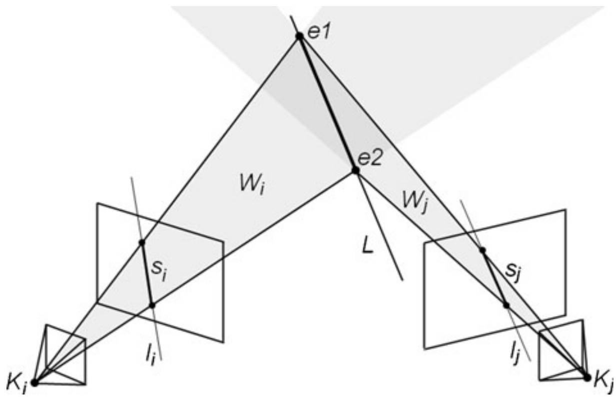
With the exception of the motion estimation process which runs on a few data on the CPU, all other stages of the pose estimation process, described in this subsection, are implemented in the GPU. This provides a brisk response time to the system, despite the large throughput of data imposed by the frequent framerate and high resolution of the three cameras. Further implementation details on the parallelization of the process can be found in [48].

3.4 Stylus-based indication

A third process combines book recognition and stylus pose-estimation to provide additional interaction. The 3D coordinates of the stylus are monitored and when the endpoint of the stylus is found, approximately, in contact with the book an additional event is triggered. This event contains the point Q of the book in contact with the stylus in the book's coordinate system, as well as the id of the corresponding page. In this way, the recipient of the event is able to find the indicated content in the book.

The following process is employed to estimate point Q. Initially, the distance of the endpoints of the stylus from the planar surface of the desk is considered. If the closest point

Fig. 7 Geometry of 3D stylus pose estimation, for two views. The detected line segments $s_{i,j}$ in images I_i and I_j define a line segment from e_1 to e_2 in 3D space, through the intersections of the planes $W_{i,j}$ that emanate from the cameras and through $s_{i,j}$ (see text for more details). When more views are available the corresponding intersections are combined to increase estimation accuracy



of the stylus is (approximately) in contact with the desk, then the point where line L intersects the plane of the desk is found, as a line-plane intersection problem. This point is then tested as to if it occurs within the spatial extend of the book at its current locus, as estimated in Sec.3.3. If Q does occur within the spatial extent of the book, then the homography corresponding to the particular page that the intersection occurs is utilized to map Q's coordinates on the electronic representation of the page. The outcome of this process is demonstrated in Fig. 8. If Q occurs outside the book no event is produced.

3.5 Handwriting and gestures recognition

This module takes as input the feedback provided by the book recognition and stylus posture and indication modules and identifies whether the user is performing a gesture or is writing text. The recognizable gestures are strongly correlated with the natural reading process (e.g., underline text, circle a word, etc.), ensuring that standard studying habits and practices are supported.

For handwriting recognition, Microsoft's Ink Analysis API [34] is used. The Microsoft Ink Analysis is an integral part of Windows 7 Tablet PC edition, and provides highly accurate handwriting and writing gestures recognition. The NHuspell spell checker [30] is also used with OpenOffice's English dictionary [38] for better recognition of the written words.

The recognized words or gestures, along with their position on the book's page and their size, are fed to the next SESIL module (Context sensitive assistance). Furthermore, the recognized page is also propagated.

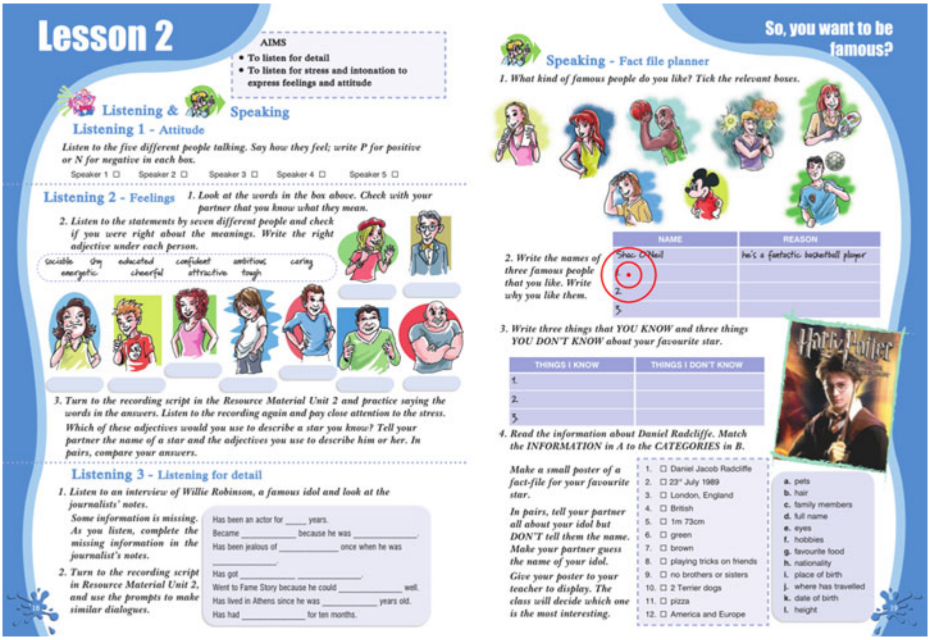


Fig. 8 Stylus-based indication of a point on the book, applied to the example of Fig. 5. The figure shows the electronic representation of the recognized page and the estimated point of stylus contact with superimposed red circles

Although Microsoft's Ink Analysis API is able to recognize a large number of gestures (about 40 different gestures), SESIL uses a reduced set, including only those gestures which can be strongly correlated with the provided educational material and exercises. Table 2 illustrates the gestures that are currently recognized by SESIL, along with the related task types.

In addition, SESIL recognizes and isolates gestures, throughout students' handwriting process, in order to be able to identify specific triggers, such as the deletion of a word that a student has written (identifying the Scratch Out gesture) or the accomplishment of a sentence (identifying the Tap gesture), which declares that a student has finished providing input and the system can therefore further process it.

3.6 Book modeling and content sensitive assistance

Every book that is included in the SESIL's recognition library has been stored in digital form (Portable Digital Format - PDF) and has also been enhanced with educational meta-data. In more detail, every page of a recognizable book is classified according to the educational material it provides (e.g., reading text, image, comprehension exercise, etc.) defining in this way the interaction and content assistance needs of the user. For example, when a user tries to accomplish a multiple choice exercise, SESIL is able to provide in real time hints and explanations about a specific section of the exercise, which is in the user's current attention area.

The classification of books' pages is contained in a XML description file for every page, in which the interactive area (hotspots) and their types are defined. Figure 9 depicts the metadata file of a SESIL's page that contains a free text field of a fill-the-gap exercise. As it

Table 2 Supported gestures by SESIL

Gesture	Exercise examples
Triangle △ Square □	• Basic geometry (e.g., draw specific geometrical shapes) • Multiple choice (e.g., square the correct answer among a number of available choices)
Circle ○	• Multiple choice (e.g., circle the correct answer among a number of available choices) • Basic geometry (e.g., draw specific geometrical shapes) • Fill-in the gap with multiple choice answers (e.g., circle the correct answer to be filled-in a sentence)
Check ✓	• Multiple choice (e.g., check with a tick the box corresponding to the correct answer among a number of available choices)
Left Right	• Word search (e.g., identify words in a matrix of letters, given a list of their definitions, by striking through the word's letters)

```
<Page id="LE_025036_11">
  <ImageSource>file:///Path/to/pages/images/025-036_Page_11.png</ImageSource>
  <HotSpots>
    <HotSpotElement>
      <BoundingPoints>
        <Point>
          <X>0.290</X>
          <Y>0.168</Y>
        </Point>
        <Point>
          <X>0.360</X>
          <Y>0.168</Y>
        </Point>
        <Point>
          <X>0.360</X>
          <Y>0.184</Y>
        </Point>
        <Point>
          <X>0.290</X>
          <Y>0.184</Y>
        </Point>
      </BoundingPoints>
      <AssetType>FREE_TEXT</AssetType>
    </HotSpotElement>
  </HotSpots>
</Page>
```


Fig. 9 SESIL page metadata information

can be observed, every page is referenced by a unique id and is accompanied by its digital image. This image can be displayed on any interactive screen near the physical book, enabling the user to interact with the hotspots of the digital form of the page. Furthermore, every page can contain a number of interactive hotspots denoting its interesting areas. Every hotspot of a page is declared by four coordination points (normalized in order to be independent of the page size), representing the four corners of the hotspot's bounding rectangle and the educational asset type that this hotspot denotes.

Every asset type is assigned a number of accepted stylus gestures (e.g., circle, underline, etc.) and/or writing text. In particular, the assigned writing text to a hotspot can be either a closed set of words or any word that is contained in the English vocabulary that is used, according to the asset type.

Anytime a page hotspot is engaged by the user's handwriting, the content sensitive assistance module evaluates that input and, according to the type of educational asset that is represented, selects the appropriate support applications and displays them on the nearby screen.

The overall decision making process for choosing the suitable applications and displaying the appropriate helping content is executed by the ClassMATE services with which SESIL interoperates.

ClassMATE aims to provide a robust and open ubiquitous computing framework suitable for a school environment that: (i) provides a context aware classroom orchestration based on information coming from the ambient environment, (ii) addresses heterogeneous interoperability of AmI services and devices, (iii), facilitates synchronous and asynchronous communication, (iv) supports user profiling and behavioral patterns discovery and (v) encapsulates content classification and supports content discovery and filtering. Complementary to ClassMATE's core an application library exists in order for other application frameworks, like SESIL, to interoperate and exploit its enhanced educative facilities.

Currently, SESIL facilitates the following handwriting – based educational assets:

Free-text input. This category includes educational assets that expect free – text input from the students, such as fill-in the gaps exercise, reading/writing exercises, etc.

Multiple choice exercises. In this category fall exercises that provide a diversity of possible answers for a specific question and the student has to select the correct answer, either by circling it, squaring it or by writing a check symbol next to it.

Match-up exercises. This type of exercises is comprised of two separate lists of elements (e.g., words, pictures, definitions, etc.) and the student has to draw a line that connects two elements, each from a different list, that are correlated in a specific manner (e.g., word – definition, word – corresponding picture, etc.)

True or false exercises. For this type of exercises students have to select if a number of statements is correct or erroneous, and in order to do so they have to make a Check gesture next to the statement in case it is right or make a Scratch out statement in case it is wrong.

Basic geometry/mathematics exercise. This type of exercise is usually encountered in teaching mathematics to elementary school students. In more detail, students are required to draw a number of specific geometric shapes (e.g., triangle, square, circle) as a response to a given mathematics problem statement.

Figure 10 illustrates a screenshot from a demo application which exemplifies how SESIL can assist a user trying to accomplish a multiple-choice exercise. In the left part of the screen the image of the active page is depicted, containing the handwritten input of the user (the circle around the word "caring"). When the user circles the word "caring" in the actual book page, SESIL recognizes the circle gesture around the hotspot of the page, which denotes that the word "caring" can be one of the user's choices. SESIL also adds the selected word under the corresponding avatar of the exercise and launches two informative applications: a) an images' collection related to the word "caring" and b) a definition, examples and synonyms of the word "caring", helping the user to strengthen his confidence regarding the correctness of his choice and allowing him to retrieve/revisit related educational information. In order

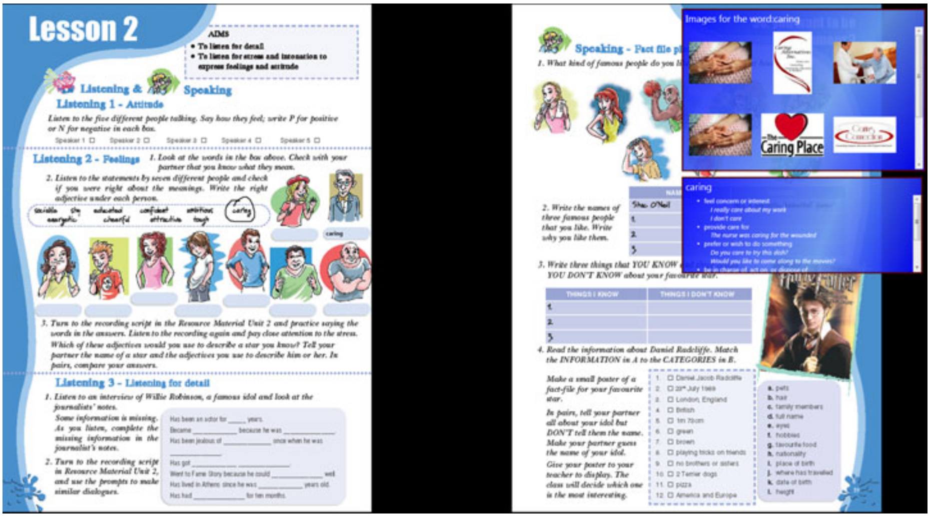


Fig. 10 SESIL screenshot example

for SESIL to decide the corresponding result to be presented to the user (e.g., in the above example, appropriate context aware help is provided when a student circles a word), it interoperates with the ClassMATE's Context Manager module that directs it to consult the User Profile module and eventually retrieve the appropriate educational assets stored in the Data Space module in the form of LOM objects [20]. In more detail, every educational content used in the classroom as a Learning Object is accompanied by LOM-based metadata that describe its educational attributes and facilitate its automatic retrieval.

In more detail, whenever SESIL receives handwriting – input from the student for a specific page's hotspot, it propagates the MIME type of interaction that is described in the ASSET_TYPE attribute of the hotspot, along with the processed handwritten input, to the ClassMATE's Application Launcher module. The Application Launcher module is part of the Context Manager and is responsible to launch the appropriate application in a valid state for its current context of use, addressing user's interaction needs according to the classroom artifact where the application has been launched. An application can be launched either directly by the classroom artifact coordination module, or indirectly by its Mime Handler delegate when handling a Mime Command fired by an application. In both cases, for an application to be launched, ClassMATE should incorporate mechanisms to both resolve from the installed applications the preferred one(s), either by name or by mime type association, and ensure that the essential security – related requirements are fulfilled. Figure 11 illustrates the overall application interaction manipulation procedure.

4 Results

The evaluation of SESIL has been performed in a setup, which employed a trinocular system installed $40cm$ above the enhanced school desk. The maximum baseline is $71.5cm$ and cameras verge at the center of the table, configured at a FOV of $43^o \times 33^o$. The main reason for spreading cameras across a baseline as wide as possible (the full extent of the desk) is robustness to occlusions and increased accuracy, as the wand is then imaged from more diverse viewpoints. Image resolution was 960×1280 pixels, except when modulated to measure effects in accuracy and computational performance. The computer hosting these cameras employed an *nVidia GeForce GTX 260* $1.2GHz$ GPU.

4.1 Book page recognition and localization

The system was 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 and by slanting the book. It was found that to retain the same confidence in page recognition the book pages cannot appear slanted more than ~60% in the image. Regarding

Fig. 11 ClassMATE's application launch mechanism for a context-aware user input

occlusions, it was found that robust page recognition requires about 30% of the pages to be visible. 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. Figure 12 demonstrates an example of successful page recognition despite a severe occlusion.

In the given setup, the error in book localization accuracy has a mean average of 5 mm (stdev: 3.7 mm), when under only mild occlusions. This accuracy is attributed again to the abundance of features that are employed in the homography computation. For the same reason, book localization accuracy starts to decrease upon occlusions that hide more than 50% of the page and, as a result, the estimated book outline exhibits jitter. Such cases are detected by the ratio of detected features, and are treated by estimating only its centroid and using a past estimate of the book's dimensions on the desktop. Nevertheless, localization accuracy under severe occlusions has still room for future improvements.

Recognition occurs in the same programming loop (one iteration corresponding to the processing of one input image) and, thus, the overall computational cost defines the achieved framerate. In the experiments, the average retrieval was below .4 sec for the evaluated books.

4.2 Stylus pose estimation

A systematic analysis of the accuracy and computational performance was conducted in order to evaluate the accuracy and framerate at which stylus pose is estimated.

A dataset annotated with ground truth was utilized to measure the pose estimation error, in a wide range of poses, consisting of 36, 360-degree yaw-rotations, in steps of 10°. The pitch angles of these rotations ranged from -70° to +80°, in steps of 10°. The dataset of about 1000 clicks was collected by mounting the stylus on a tripod with 2° of freedom (pitch, yaw) and marked rotation gratings. To study the effects of resolution and number of views in the accuracy of estimates, they were modulated as shown in Table 3. It can be concluded that the method is sufficiently accurate for purposes of indicating points on the desktop and that accuracy gracefully degrades to the reduction of input data.

The localization accuracy of the stylus point of contact was always below .5 cm, thereby facilitating the robust stylus-based indication of paragraphs, figures, equations, etc., within the book.

To evaluate the computational performance of the method, performance was measured while modulating the number of views and resolution, using 2 and 3 cameras. Table 4 reports the

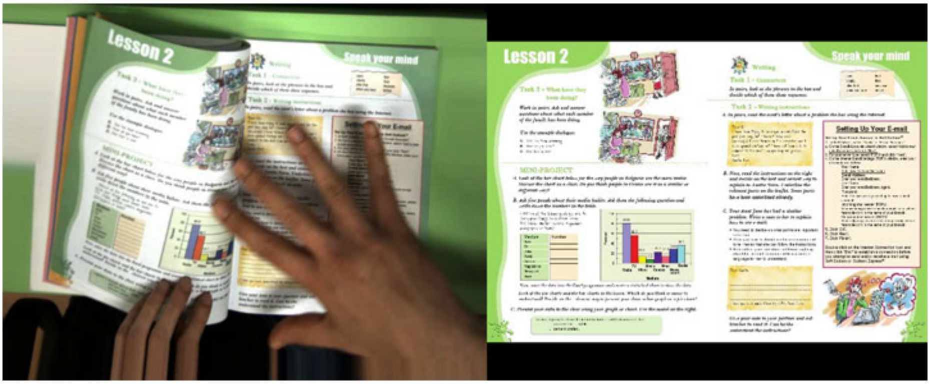


Fig. 12 Book page recognition under occlusions. Due to the representation of page appearance using local features, recognition is possible despite severe occlusions

Table 3 Mean error and standard deviation results, in degrees, obtained in the accuracy evaluation of stylus pose estimation

	480×640		960×1280	
Views	Yaw	Pitch	Yaw	Pitch
2	2.0 (3.1)	1.0 (1.4)	1.4 (4.1)	0.8 (1.9)
3	1.2 (1.1)	0.7 (0.7)	1.2 (1.1)	0.7 (0.7)

achieved framerate of the system with respect to the amount of input data, that is, the number of cameras and their resolution; as shown above the more the input data, the greater the accuracy of the system. It is observed that the method is efficiently parallelized in the GPU and that it scales well with the amount of input, yielding also a brisk response that can support interactivity.

4.3 Usability evaluation

Following a user-centred design approach, and since a functional yet early prototype of SESIL was available, a formative usability evaluation was carried out by usability and interaction design experts, aiming to assess the usability and usefulness of the system. The evaluation was carried out by four experts, who were asked to carry out a heuristic evaluation in order to identify potential usability problems. Heuristic evaluation is an informal usability inspection method [37] and involves having a small number (ideally three to five) usability specialists judge whether each dialogue element follows established usability principles (the "heuristics"). The output of the process is a list of the usability problems in the interface with references to those usability principles that were violated by the design in each case. Once the usability problems have been reported individually by each evaluator, an evaluation facilitator creates an overall report. Then, each evaluator is asked to rate the complete list or problems according to their severity, and an average score is calculated for each problem in order to be able to prioritize their elimination. Severity ratings range from 0 (I don't agree this is a usability problem at all) to 4 (usability catastrophe).

After interacting with the system, the evaluators were asked to fill-in a short questionnaire in order to assess conformance with the five dimensions of usability [40] and the anticipated overall user experience [22, 24]. The five dimensions of usability constitute an evolvement of the traditional ISO 9241 characteristics of usability, aiming to create a more precise description both the goals for, and experience of, using a product. Namely, the five dimensions of usability are: effective, efficient, engaging, error tolerant, and easy to learn. In addition to the traditional notion of usability, user experience highlights non-utilitarian aspects of interactions, shifting the focus to user experience. It is associated with abroad range of fuzzy and dynamic concepts, including emotional, affective, experiential, hedonic, and aesthetic variables; typical examples of user experience attributes include fun, pleasure, pride, joy, surprise, and intimacy, which however are only a subset of a growing list of human values. As a result, given the nature of the system being evaluated, it was considered necessary not only to assess its usability with a traditional usability evaluation method, but also to assess the overall user experience as anticipated by the experts.

The heuristic evaluation resulted in 7 usability improvement suggestions, 2 of which were classified as major problems, 3 as minor problems and 2 as aesthetic problems only. The

Table 4 Achieved framerate (temporal frequency of operation) for different number of views and image resolutions

Views	480×640	960×1280
2	30 Hz	15 Hz
3	30 Hz	10 Hz

evaluation suggested that the dictionary and images application could be enhanced with further functionality, allowing for example users to add a word to their personal vocabulary.

The questionnaire given to experts in order to assess the usability and overall anticipated user experience comprised of 13 questions, providing pairs of contrasting characterizations for the system. The answers were analyzed individually (see Fig. 13), where the rating for each attribute ranges from 1 to 7 (1 represents the most positive attitude, e.g., pleasant, and 7 the most negative attitude, e.g., unpleasant). Answers were also grouped in three categories, namely usability, learnability and user experience, with each one of the categories rated on a scale from 0 to 10. Usability received an average score of 8.21 (stdev: 0.28), learnability scored 8.09 in average (stdev: 0.3) and user experience was rated with 7.5 (stdev: 0.27). These results can only be used as qualitative measures, while more detailed quantitative results will be obtained in future evaluations with users.

5 Conclusion

This paper has presented SESIL, a system targeted to enhance reading and writing activities on physical books through unobtrusive monitoring of users' gestures and handwriting and the display of information related to the current users' focus of attention. SESIL innovatively

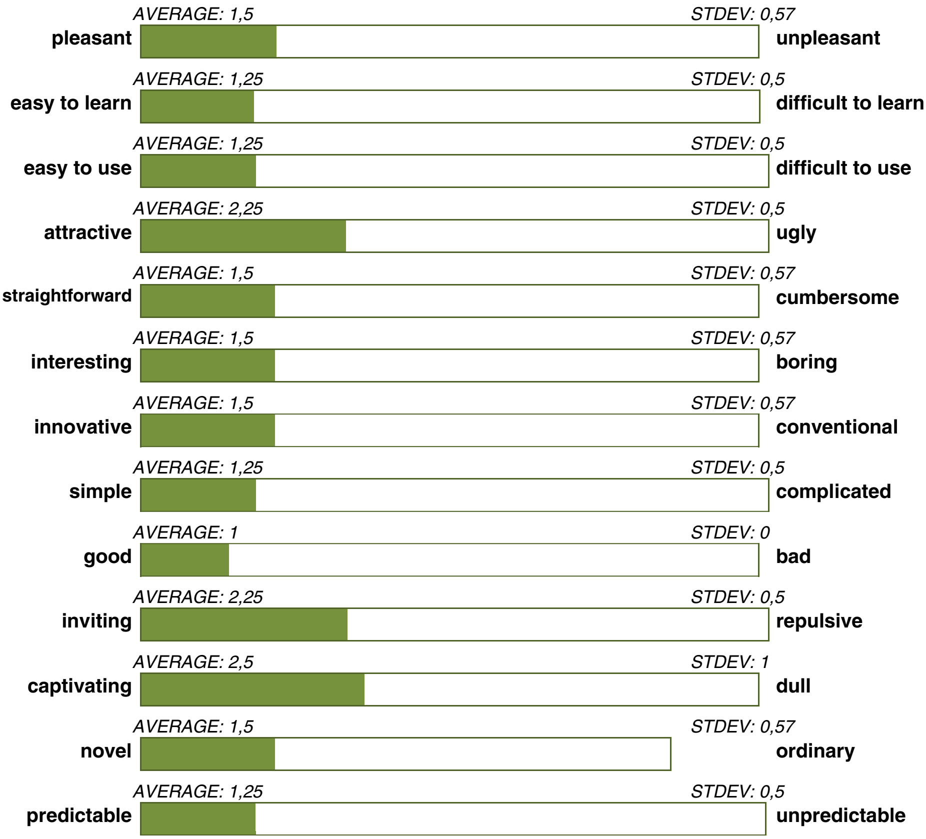


Fig. 13 Questionnaires results

combines book and page recognition, stylus position identification, and handwriting recognition without requiring any special device apart from cameras. Additionally, SESIL exploits educational metadata on the book's content to decide at run-time the type of additional information and support need to be provided in a context-dependent fashion.

SESIL seamlessly combines reading and handwriting on physical books with the provision of related digital content through natural interaction. The performed tests have confirmed that SESIL is robust and reliable enough for practical use, and formative usability evaluation has yielded positive results. Future work will focus on user-based evaluation in the context of a (simulated) AmI classroom environment, also assessing the educational benefits and implications of the overall approach.

Acknowledgments This work is supported by the Foundation for Research and Technology Hellas – Institute of Computer Science (FORTH – ICS) internal RTD Programme 'Ambient Intelligence and Smart Environments' [15]. The authors would like to thank Mrs. Stavroula Ntoa for her contribution to the usability evaluation of SESIL.

References

1. Aarts E, Encarnacao JL (2008) True visions. The Emergence of Ambient Intelligence. Springer, ISBN 978-3-540-28972-2
2. Abrami PC, Bernard RM, Wade CA, Schmid RF, Borokhovski E, Tamim R, Surkes M, Lowerison G, Zhang D, Nicolaidou I, Newman S, Wozney L, Peretiatkowicz A (2006) A review of E-learning in Canada: a rough sketch of the evidence, gaps and promising directions. Canadian Journal of Learning and Technology, 32 (3). http://www.cjlt.ca/index.php/cjlt/article/view/27/25. Accessed 16 February 2011
3. Anoto, Development Guide for Service Enabled by Anoto Functionality. 2002
4. Antona M, Margetis G, Ntoa S, Leonidis A, Korozi M, Paparoulis G, Stephanidis C (2010) Ambient Intelligence in the classroom: an augmented school desk. W. Karwowski & G. Salvendy (Eds.), Proceedings of the 2010 AHFE International Conference (3rd International Conference on Applied Human Factors and Ergonomics), Miami, Florida, USA, 17-20 July
5. Artifex Software, Inc. MuPDF: a lightweight PDF and XPS viewer. http://www.mupdf.com/. Accessed 16 February 2011
6. Ayache N, Lustman F (1987) Fast and reliable passive trinocular stereovision. ICCV, pp. 422–427
7. Baillard C, Schmid C, Zisserman A, Fitzgibbon A (1999) Automatic line matching and 3D reconstruction of buildings from multiple views. ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery
8. Billinghurst M, Kato H, Poupyrev I (2001) The MagicBook: moving seamlessly between reality and virtuality. IEEE Computer Graphics, pp. 6–8
9. Bookstein FL (1991) Morphometric tools for landmark data. Cambridge University Press
10. Cook DJ, Augusto JC, Jakkula VR (2009) Ambient intelligence: technologies, applications, and opportunities. Pervasive Mobile Comput 5(4):277–298
11. Cook DJ, Das SK (2007) How smart are our environments? An updated look at the state of the art. J Pervasive Mobile Comput 3(2):53–73
12. Duda RO, Hart PE (1972) Use of the hough transformation to detect lines and curves in pictures. Comm ACM 15:11–15
13. Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Comm ACM 24:381–395
14. Forsberg AS, LaViola Jr. JJ, Zeleznik RC (1998) ErgoDesk: a framework for two- and three-dimensional interaction at the ActiveDesk. Second International Immersive Projection Technology Workshop, pp. 11–12
15. FORTH-ICS Ambient Intelligence Programme. http://www.ics.forth.gr/ami. Accessed 16 February 2011
16. Gelb A (1974) Applied optimal estimation. MIT Press
17. Grasset R, Duenser A, Seichter H, Billinghurst M (2007) The mixed reality book: a new multimedia reading experience. CHI '07 extended abstracts on Human factors in computing systems, pp. 1953–1958
18. Hartley RI, Zisserman A (2004) Multiple view geometry in computer vision
19. Hile H, Kim J, Borriello G (2004) Microbiology tray and pipette tracking as a proactive tangible user interface. Pervasive Computing, pp. 323–339

20. IEEE LOM (2002) Draft standard for learning object metadata. IEEE Learning Technology Standards Committee, IEEE 1484.12.1
21. Ishii H, Ullmer B (1997) Tangible bits: towards seamless interfaces between people, bits and atoms. CHI. pp. 234–241
22. ISO FDIS 9241-210:2009 Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems (formerly known as 13407). International Organization for Standardization (ISO). Switzerland
23. Kobayashi M, Koike H (1998) EnhancedDesk: integrating paper documents and digital documents. Asia Pacific Computer Human Interaction (APCHI'98). IEEE CS, pp. 57–62
24. Law E, Roto V, Hassenzahl M, Vermeeren A, Kort J (2009) Understanding, scoping and defining user experience: a survey approach. Proceedings of the Human Factors in Computing Systems conference, CHI'09, April 4-9, 2009, Boston, MA, USA
25. Leonidis A, Margetis G, Antona M, Stephanidis C (2011) ClassMATE: enabling ambient intelligence in the classroom. World Academy of Science, Engineering and Technology, 66: 594 – 598. http://www.waset.org/journals/waset/v66/v66-96.pdf. Accessed 16 February 2011
26. Liao C, Guimbreti F, Hinckley K, Hollan J (2008) Papiercraft: a gesture-based command system for interactive paper. ACM Transactions in Computer-Human Interactions, 14 (4), Article 18, 27 pages
27. Lowe DG (1987) 3D object recognition from single 2D images. Artif Intell 3:355–397
28. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110
29. Luff P, Heath C, Norrie M, Signer B, Herdman P (2004) Only touching the surface: creating affinities between digital content and paper. In Proceedings of the 2004 ACM conference on Computer supported cooperative work (CSCW '04). ACM, New York, NY, USA, 523–532
30. Maier hofer software, NHunspell: C#/.NET free Spell Checker, Hyphenator, Thesaurus. Hunspell spelling corrector and hyphenation for C# and Visual Basic. http://nhunspell.sourceforge.net/. Accessed 16 February 2011
31. Mangen A, Velay J-L (2009) Digitizing literacy: reflections on the haptics of writing. Zadeh MH (Ed.), Advances in Haptics, pp. 385–401
32. Martinec D, Pajdla T (2003) Line reconstruction from many perspective images by factorization. CVPR, pp. 497–502
33. Michel D, Argyros AA, Grammenos D, Zabulis X, Sarmis T (2009) Building a multi-touch display based on computer vision techniques. IAPR Conference on Machine Vision Applications, May 20-22. Hiyoshi Campus, Keio University, Japan
34. Microsoft, Ink Analysis Overview (Windows). http://msdn.microsoft.com/en-us/library/ms704040(v=VS.85).aspx. Accessed 16 February 2011
35. Moons T, Frere D, Vandekerckhove J, Gool L (1998) Automatic modeling and 3D reconstruction of urban house roofs from high resolution aerial imagery. ECCV, pp. 410–425
36. Nickel K, Stiefelhagen R (2007) Visual recognition of pointing gestures for human-robot interaction. Image Vision Comput 25:1875–1884
37. Nielsen J, Mack RL (1994) Usability inspection methods. John Wiley & Sons Inc., New York, pp 25–61
38. Open Office Wiki, Dictionaries. http://wiki.services.openoffice.org/wiki/Dictionaries. Accessed 16 February 2011
39. Quan L, Kanade T (1997) A_ne structure from line correspondences with uncalibrated a_ne cameras. PAMI 19:834–845
40. Quensbery W (2003) The five dimensions of usability. Albers MJ, Mazur B (Eds.) Content and Complexity, Routledge, pp. 75–94
41. Schmalstieg D, Fuhrmann V, Hesina G, Szalav Z, Encarna LM, Gervautz M, Purgathofer W (2002) The studierstube augmented reality project. Presence: Teleoper. Virtual Environ 11:33–54
42. Shi Y, Xie W, Xu G, Shi R, Chen E, Mao Y, Liu F (2003) The smart classroom: merging technologies for seamless tele-education. IEEE Pervasive Computing (April–June), pp. 47–55
43. Simon A, Dressler A, Kruger H, Scholz S, Wind J (2005) Interaction and co-located collaboration in large projection-based virtual environments. IFIP Conference on Human-Computer Interaction, pp. 364–376
44. Smith R, Chang S (1996) VisualSEEk: a fully automated content-based image query system. ADM Multimedia, pp. 87–89
45. Torr PHS, Zisserman A (1997) Robust parameterization and computation of the trifocal tensor. Image Vision Comput 15:591–605
46. Wellner P (1993) Interacting with paper on the DigitalDesk. Commun. ACM 36, 7 July, pp. 87–96
47. Woo D, Park D, Han S (2009) Extraction of 3D line segment using disparity map. Digital Image Processing, pp. 127–131
48. Zabulis X, Koutlemanis P, Baltzakis H, Grammenos D (2011) Multiview 3D pose estimation of a wand for human-computer interaction. International Symposium on Visual Computing, September 26-28, Las Vegas, Nevada, USA

49. Zabulis X, Sarmis T, Tzevanidis K, Koutlemanis P, Grammenos D, Argyros AA (2010) A platform for monitoring aspects of human presence in real-time. International Symposium on Visual Computing, Las Vegas, Nevada, USA, November 29 - December 1

George Margetis holds a degree in Computer Science and M.Sc. in “Computer Networks and Digital Communications” and “Information Systems”. He is a member of the Human-Computer Interaction Laboratory of FORTH-ICS since 2005. His past work includes network traffic measurement and analysis in high-speed networks, resource control and service differentiation in wired networks. His current work focuses on Universal Access, Design for All, eLearning, interaction design, and Ambient Intelligence. He has participated in a number of European research projects, including ASK-IT, MICOLE, eUser, etc. He has also participated in several national projects, including eGovernment Services Portal for Disabled, Educational Portal for the Region of Crete, the portal of the World Hellenic Inter-Parliamentary Association, and the portal of the Hellenic Parliament Foundation, where he was technical coordinator.

Xenophon Zabulis is currently a Researcher at the Institute of Computer Science – Foundation for Research and Technology, Hellas. He received the B.A., M.S. and Ph.D. degrees in Computer Science from the University of Crete, Greece, in 1996, 1998 and 2001, respectively. From 2001 until 2003 he has worked as a Postdoctoral Fellow at the Computer and Information Science Department, at the interdisciplinary General Robotics, Automation, Sensing and Perception laboratory and at the Institute for Research in Cognitive Science, both at the University of Pennsylvania, USA. In addition, during 2004 to 2007, he has worked as a Research Fellow at the Institute of Informatics and Telematics, Centre of Research and Technology Hellas, Greece. His research interests include stereo & multiple-view vision, real-time 3D reconstruction of static & dynamic scenes, person localization and head pose estimation, camera networks, and applications of computer vision in interactive environments.

Panagiotis Koutlemanis was born in Greece, in 1980. He received his B.Sc. degree in Music Technology and Acoustics from the Technological Educational Institute of Crete, with a major in virtual reality using binaural audio. He has worked as a developer for the Technological Educational Institute of Crete from 2005 to 2008. Since 2009, he has been working as a developer at the Institute of Computer Science - Foundation for Research and Technology, Hellas (FORTH), participating in various research programs.

Margherita Antona holds a BA in Computational Linguistics from the University of Pisa, Italy, and a PhD in Human-Computer Interaction from the University of Kent, UK. She is member of the HCI Laboratory of ICS-FORTH since 1993. Since 2004, she is Coordinator of the Centre for Universal Access & Assistive Technologies of ICS-FORTH, and since 2009 Coordinator of the AmI Classroom activity of the ICS-FORTH AmI Programme. Her research interests include universal access, design for all, computer-supported user interface design, adaptive and intelligent interfaces, assistive technologies, eLearning and Ambient Intelligence. She has participated in more than 20 European and national R&D projects, and is the scientific responsible for the participation of ICS-FORTH in the projects AAL-REMOTE and FP7-ICT-VERITAS. She has coauthored more than 60 scientific publications. She is member of Program Committee and Paper Review Committee in various international conferences and workshops, including UAHCI, AVI, ECDL, MobileHCI, IADIS WWW/Internet, ICSI HCI/MIS, W4A, WEBIST, PETRA, as well as Program Co-chair of UAHCI 2013. She is also a member of the Editorial Board of the Universal Access in the Information Society International Journal. From 1988 to 1992 she was researcher at the Institute of Computational Linguistics of CNR in Pisa.

Constantine Stephanidis Professor at the Department of Computer Science of the University of Crete, is the Director of the Institute of Computer Science (ICS), Foundation for Research and Technology – Hellas (FORTH), Head of the Human - Computer Interaction Laboratory, and of the Centre for Universal Access and Assistive Technologies, and Head of the Ambient Intelligence Programme of ICS-FORTH.

In the beginning of the '90s he introduced the concept and principles of Design for All in Human-Computer Interaction and for Universal Access in the evolving Information Society. Over the past 25 years, he has been engaged as the Scientific Responsible in more than 50 National and European Commission funded projects in the fields of Human-Computer Interaction, Universal Access and Assistive Technologies.

Currently, he is coordinating multidisciplinary activities aiming to conceptualize, design and develop in vitro (and deploy in vivo) pioneering, innovative technologies and applications for Smart Environments, capable of "understanding" and proactively addressing individual human needs, following a human-centred approach.

He has published more than 550 technical papers in scientific archival journals, proceedings of international conferences and workshops related to his fields of expertise.

In 2010, Prof. Stephanidis was elected member of the Informatics Section of the Academia Europaea, a functioning European Academy of Humanities, Letters and Sciences.

http://www.ics.forth.gr/index_main.php?l=e&c=72