

MIRELE: A MIXED-REALITY SYSTEM FOR LEARNING A TASK DOMAIN

Zardosht Hodaie, Sajjad Taheri and Bernd Brügge

*Technical University of Munich. Faculty of Informatics. Chair for applied Software Engineering
Boltzmannstr. 3, 85748 Garching b. München, Germany*

ABSTRACT

MIRELE is an interactive tabletop for situated domain learning that helps students learn names of the physical objects and their relationships by projecting the names, relationships, and additional information onto the real objects. It is build based on camera-projector system and offers a domain-independent fast authoring system. Using the authoring system, the teachers can create the learning materials themselves without any programming. MIRELE is designed for domain-independence, unobtrusiveness and usability for manual activities, and offers a simple low-cost solution that can be utilized very fast in the classroom.

KEYWORDS

Interactive Tabletops, Situated Vocabulary Learning, AR Authoring

1. INTRODUCTION

An essential part of learning any domain is learning its vocabulary, concepts, and the relationships between them. In this paper we focus on the domain of manual-procedural activities (MPAs) (Hodaie, Haladjian & Bruegge 2018). Manual-procedural activities involve manipulating real physical objects by following a given procedure. Examples of MPAs are numerous, such as repairing a motor, assembling a part, and different crafts. In all these activities we a) manipulate physical objects and materials either directly by hands, or indirectly using tools; and b) follow steps of a given procedure, either by looking it up in the instruction manual, or knowing it by heart, or imitating an instructor. As part of learning a manual-procedural activity, a student must learn the names of the objects, materials, and tools, and how they relate to each other. Acquisition of this domain knowledge is part of every Vocational Education and Training program (Rauner et al. 2012).

A motivation behind our work is the current refugee crisis facing the European countries. Many of the refugees are skilled professionals that could be integrated into the local economies. However, a challenge towards incorporating them quickly into the job market is the language barrier. These professionals already know their corresponding domain of work, but they don't know the vocabulary in the language of the host country. So our system aims to aid them in faster learning of their technical vocabulary.

In this paper we propose a mixed-reality system based on a camera-projector setup to assist vocabulary and concept learning. Our assumption is that a mixed-reality system that involves the real objects and provides a situated (Lave & Wenger 1991) and embodied (Abrahamson & Lindgren 2014) learning experience will be more effective in learning and retention of the vocabulary and concepts of a domain.

The system consists of three main components: an interactive tabletop that recognizes and tracks real objects and projects names and different additional information onto them; A domain-independent authoring system that allows teachers to create learning content without any programming; A simple GUI for creating object recognition components.

2. RELATED WORK

The Bloom's Taxonomy of learning objectives considers acquiring knowledge as the basis of all other cognitive processes of learning (Krathwohl 2002). The knowledge itself is further categorized into factual knowledge (e.g. terminology), conceptual knowledge (e.g. categories and interrelations between concepts), and procedural knowledge (e.g. methods and procedures for doing something).

There is evidence in the literature that affordances such as embodiment and interactivity provided by Augmented Reality (AR) and tangible user interfaces improve learning and knowledge acquisition (Schneider et al. 2010, Diegmann et al. 2015). Manches et al. (2009) discuss the perceptual and manipulative properties of physical objects and how these properties can support learning for children. Working with physical objects and hands-on learning is also an integrated part of engineering disciplines (Carlson & Sullivan 1999). Accordingly, we argue that when the domain of learning is the physical world, such as tools and objects required for accomplishing a physical task, interacting with the real objects will benefit the learning performance.

Another related area of research for our work is the situated vocabulary learning. Several systems have utilized the pedagogical theories of hands-on, authentic, and situated learning (Ozverir & Herrington 2011) in the field of vocabulary learning. Ogata et al. (2004) use RFID tags on objects and environment attached sensors in order to track user's environment and present the relevant vocabulary and phrases. Santos et al. (2016) created a marker-based mobile vocabulary learning app and showed a slight improvement in learning gain as well as a significant positive attitude of the users toward learning. Vazquez et al. (2017) created a mixed-reality app using Microsoft HoloLens that uses cloud-based object recognition to enable a serendipitous language learning experience with dynamic content.

Our approach is distinguished from related work in two aspects. We offer a fast and domain-independent authoring system that allows flexible definition of learning content directly by teachers. Additionally, we offer a simple GUI for fast creation of object recognition components that can be used in the system. As a result, our system is not limited to hard-coded pre-defined content, and can be used directly by teachers of different domains.

3. MIRELE

Figure 1 (left) shows use cases of MIRELE. The system has two actors: the teacher and the student. Teachers can define learning content in the Authoring mode. This learning content is presented in Training mode to the student. The teacher can also define exercises that are presented to the student in Exercise mode.

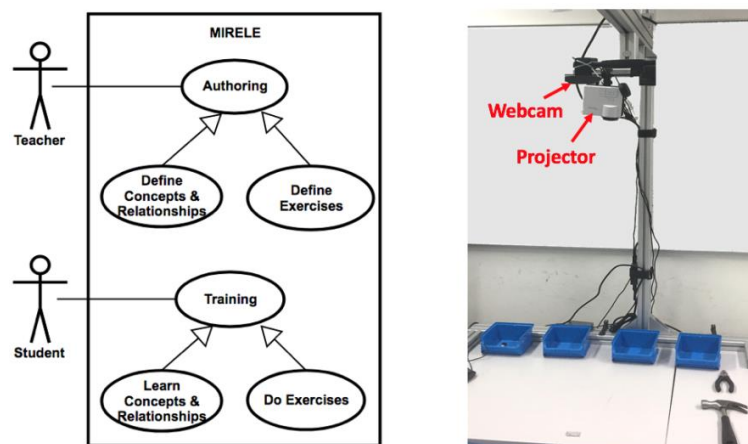


Figure 1. Use case of MIRELE and the camera-projector setup

We considered the following design goals in creating MIRELE: Low-cost, easy to setup, and easy-to-use system; and fast and domain-independent authoring. Accordingly, we chose a camera-projector setup, because the required equipment can easily be acquired with comparatively low price in the consumer market. Furthermore, this setup allows for easy hands-free interaction without the need for bulky and tethered head-mounted displays. The same setup is used by the teachers to create the learning content quickly using a graphical user interface. Figure 1 (right) shows our camera- projector setup.

MIRELE has two modes: Authoring and Training. In Authoring mode, the teacher interacts with the camera-projector setup and authoring GUI on a PC to create learning contents. This step includes defining objects and their annotations and defining exercises. The annotations can be different multi-media content like text, arrows, video, and audio. In Training mode, the student interacts with the tabletop by putting and manipulating objects on the table, or selecting the objects and annotations using a tracked selection stick. In Training mode, the content defined by the teacher is projected onto the tabletop.

Learning content are presented as AR annotations projected onto the objects on the tabletop. This include textual annotations, arrows, circles, checkboxes, images, audio, and video snippets. Each annotation can be either attached to an object or placed on a fixed position on the tabletop. Furthermore, the teacher can define named relationships between objects. In Training mode, the relationships are shown as straight labeled lines between objects. For example, different kinds of pliers can be grouped together using relationship lines.

The system does not consider the semantics of the content. The teacher defines what are the learning content and their semantics. This can be names of the materials, objects, and tools, names of the parts of an object, and different relationships between objects and which objects are required for a task. Similarly, exercises are defined as membership questions. For example, "Which objects are needed to accomplish task X?". The system considers the content merely as different kinds of AR annotations.

In order to dynamically project annotations onto the physical objects, the table needs to know which objects are on the table, where are their positions, and what are their orientations. This is achieved using object recognition components. For each set of objects in a domain, that are involved in a learning session, the teacher needs first to create and include the required object detectors. An object detector is either based on a machine learning model (marker-less object detection) or based on markers. In the case of marker less object detection, we offer a simple UI in which the teacher selects objects that need to be detected. The program then creates an object detection package that include a trained ML model. Depending on the hardware used, this process may take a long time, up to several hours. For marker-based object detection, the teacher associates physical objects with the markers, by selecting them in the UI and assigning a marker ID to them. In addition to object detection, the system also offers detection of a selection stick, using which the student can interact with objects and annotation. For example, selecting an object can trigger an audio annotation that names the object. Or selecting a video annotation will start playing it.



Figure 2. The system screenshots (see text for descriptions)

Figure 2 shows the screenshots of the systems in different modes. In Authoring mode (Figure 2a), the teacher sees a top-view of the table provided by the camera and graphically interacts with it to define the content. The teacher puts the objects on the table and selects them to be tracked by the systems. He then adds different kinds of annotations like textual descriptions, arrows, or video snippets attached to the objects or at fixed positions on the scene. In the Training mode (Figure 2b), the student puts the objects on the table, the system detects the objects and projects the annotations provide by the teacher onto them. The student can select objects and annotation, for example selecting a video annotation that demonstrates how to use a tool. The system also projects the relationships between objects as labeled straight lines. The teacher defines exercises as questions that have objects as answer (Figure 2c). The exercise consists of a question text and a set of objects as the correct answer. For example, "Which tool do you use to pull out a nail?" and for the correct answer the student should put the "pincer" on the table. It is also possible to define an exercise as a multiple-choice question using checkboxes, or as a selection task, in which the student should select the right

object. In exercise mode (Figure 2d), the system projects the question on the table. The student answers the question by putting objects on the table. The system detects the objects put on the table and gives feedback on correctness of the answer.

4. CONCLUSION AND FUTURE WORK

In this paper we introduced our ongoing work on MIRELE, an interactive tabletop intended to support learning of a task domain in manual-procedural activities. MIRELE offers a simple authoring tool for the interactive tabletop, using which the teachers themselves can create learning content, without depending on programmers. We believe this low-cost and simple system will lower the entrance barrier for teachers to experiment with new possibilities of augmented reality and tangible interaction in the classroom. Furthermore, the authoring system together with the domain-independence focus of MIRELE allows rapid creation of learning materials for different domains. We believe this opens new opportunities for research on the use of augmented reality in the classroom, that is conducted directly by pedagogical experts themselves and in which the cost and effort of content creation is significantly reduced.

We are currently conducting a pilot between-subject user study to evaluate the effectiveness of MIRELE for learning the vocabulary of simple task domain consisting of 10 tools from a typical house-hold toolbox in German language. Our assumption is that the participant who learn with physical objects will perform better in terms of retention of the object names compared to those who use a paper-based learning material. In the next step, we plan to conduct usability studies with teachers of a vocational school, in order to evaluate how they use MIRELE and get feedback for improving its feature set.

REFERENCES

- Abrahamson, D. and Lindgren, R., 2014. *Embodiment and embodied design*. The Cambridge handbook of the learning sciences, 2, pp.358-376.
- Carlson, L.E. and Sullivan, J.F., 1999. Hands-on engineering: learning by doing in the integrated teaching and learning program. *International Journal of Engineering Education*, 15(1), pp. 20-31.
- Diegmann P., Schmidt-Kraepelin M., van den Eynden S., Basten D., 2015. Benefits of augmented reality in educational environments—A systematic literature review, In *Proc. 12th Int. Conf. Wirtschaftsinformatik*, pp. 1542-1556.
- Hodaie, Z., Haladjian, J. and Bruegge, B., 2018, June. TUMA: Towards an Intelligent Tutoring System for Manual Procedural Activities. In *International Conference on Intelligent Tutoring Systems* (pp. 326-331). Springer.
- Krathwohl, D.R., 2002. A revision of Bloom's taxonomy: An overview. *Revising Bloom's Taxonomy*. Theory into practice, 41(4), pp. 212-218. The H.W. Wilson Company.
- Lave, J. and Wenger, E., 1991. *Situated learning: Legitimate peripheral participation*. Cambridge university press.
- Manches A., O'Malley C., Benford S. 2009. Physical manipulation: evaluating the potential for tangible designs. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*. ACM, 77–84.
- Ogata, H., Akamatsu, R. and Yano, Y., 2004. Computer supported ubiquitous learning environment for vocabulary learning using RFID tags. In *IFIP World Computer Congress*, TC 3 (pp. 121-130). Springer, Boston, MA.
- Ozverir, I. and Herrington, J., 2011, June. Authentic activities in language learning: Bringing real world relevance to classroom activities. In *EdMedia+ Innovate Learning* (pp. 1423-1428). Association for the Advancement of Computing in Education (AACE).
- Rauner, F., Heinemann, L., Maurer, A. Haasler, B., 2012. *Competence Development and Assessment in TVET (COMET): Theoretical framework and empirical results* (Vol. 16). Springer Science & Business Media.
- Santos, M.E.C., Taketomi, T., Yamamoto, G., Rodrigo M., and Kato, H., 2016. Augmented reality as multimedia: the case for situated vocabulary learning. *Research and Practice in Technology Enhanced Learning*, 11(1), p.4.
- Schneider, B., Jermann, P., Zufferey, G. and Dillenbourg, P., 2010. Benefits of a tangible interface for collaborative learning and interaction. *IEEE Transactions on Learning Technologies*, 4(3), pp. 222-232.
- Vazquez, C.D., Nyati, A.A., Luh, A., Fu, M., Aikawa, T. and Maes, P., 2017, May. Serendipitous language learning in mixed reality. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2172-2179). ACM.