Toys++

AR embodied agents as tools to learn by building

Luca Simeone, Salvatore Iaconesi
FakePress, University La Sapienza
Rome, Italy

Post-print version

Citation: Simeone, L., Iaconesi, S. (2010) Toys++: Augmented reality embodied agents as tools to learn by building. Proceedings of IEEE ICALT, Sousse (Tunisia), 5-7 July 2010

Abstract—**This paper presents a prototype for an augmented-reality based toy. Toys++ is grounded on the concept that the actual activity of building tangible artifacts can speed up learning processes. Toys ++ aims at assembling a framework that will allow the use of existing physical components of the toy as triggers. When the toy is placed under the webcam, a pre-trained 3D feature recognition system scans the entire figure, trying to identify some specific components. If any of these elements are recognized, the system will retrieve and show educational content from selected sources (texts, videos, pictures).**

Keywords-component; augmented reality; enhanced learning; constructionism; mathetics

I. Introduction

Augmented reality technologies define the possibility of reconsidering the ways in which we communicate, interact, relate, behave, including the ways in which we exchange, distribute, share, disseminate knowledge and information. In this scenario the ideas of "learning", "teaching" and "communicating" have been reinterpreted by extending the spaces and tools that can be used in these practices.

Toys++ merges the characteristics of real-life toys and web-based educational tools creating hybrid devices that work as tangible learning objects. Within this framework, new learning grammars, uses and strategies can be fostered, especially ones closely related to constructionist paradigms.

Constructionism is based on the idea of active construction. Kids (and adults) seem to learn quickly and deeper if they are involved in activities where they are asked to actively construct new knowledge, rather than just having chunks of information poured into their minds.

Constructionist theories are also grounded on the concept that the actual activity of building tangible artifacts can speed up learning processes.

"Constructionism — the N word as opposed to the V word — shares contructivism's view of learning as "building knowledge structures" through progressive internalization of actions… It then adds the idea that this happens especially felicitously in a context where the learner is consciously engaged in constructing a public entity, whether it's a sand castle on the beach or a theory of the universe" (Papert 1991:1) [1].

Following Ackermann's words (2001:2):

“Papert draws our attention to the fact that “diving into” situations rather than looking at them from a distance, that connectedness rather than separation, are powerful means of gaining understanding. Becoming one with the phenomenon under study is, in his view, a key to learning” [2].

In this sense toys, especially ones that engage the kids in construction activities like Lego, Playmobil, or wooden blocks, can be considered wonderful devices for learning. Our idea is to get these physical artifacts somewhat connected to some information retrieval systems that can show educational content related to the toy. If the child is building a Lego historical vessel, she may also be interested in knowing something more about the dynamics behind the physics of flotation, or on sailing techniques, or on traditional sea routes that were connecting the sea towns closer to where she actually lives.

Our challenge was therefore to design a framework where physical components of toys could be read through a sensor (e.g. a webcam), processed and recognized through 3D pattern recognition systems, thus acting as triggers to retrieve and show educational content (e.g. on the screen of the computer).

One of our main design constraints was to try to use existing technologies, already available at the mass market level, in order to create an open framework that could be virtually replicated at any house equipped with a basic PC and a webcam.

II. The Toys++ Framework

Augmented reality (AR) is a popular topic: extensive scientific literature and many design concepts have already been produced.

Some toy manufacturers have launched several AR kits on the market [3]. Most of them run on a PC environment and they are mostly based on physical tags that can be read from the webcam to act as triggers retrieving special content shown on the screen.

Our idea is to take the concept of augmented-reality toys to another level. Toys ++ aims at assembling a framework that will allow the use of existing physical components of the toy as triggers.

We built an initial prototype of the system using a Processing library that runs on an iMac equipped with the internal webcam. We tested it using a common airplane plastic model. When the airplane model is placed under the webcam, a pre-trained 3D feature recognition system scans the entire figure, trying to identify some specific components (the engines, the wings, the landing gear, the cockpit). If any of these elements are recognized, the system shows additional educational content on the screen of the iMac (text, video, music).

The educational content is mainly targeted to 5-14 year-old kids and is retrieved from selected online sources: National Geographic Kids, Smithsonian National Air and Space Museum, and the Exploratorium in San Francisco. The content ranges from historical images and videos, to special lessons in physics, air dynamics, quizzes, webgames, downloadables (origami, activities sheets).

Additional content for older kids has been collected from sources such as Wikipedia and YouTube (e.g. videos explaining how wings work from the Department of Engineering at the University of Cambridge), so that kids or parents can choose suitable educational content for their specific age and level of knowledge.

Another interesting feature of Toys++ is the AR instruction manual. The airplane model is usually packaged and distributed in boxes where kids can find all of the different pieces that have to be connected or glued together in order to build the actual airplane. A paper set of instructions comes inside the box.

Toys++ allows the creation of digital, interactive instruction manuals. The system has been trained to identify specific pieces of the model and to show interactive instruction on the screen on-the-fly, while the user is performing that specific activity (and therefore when she really needs that specific help).

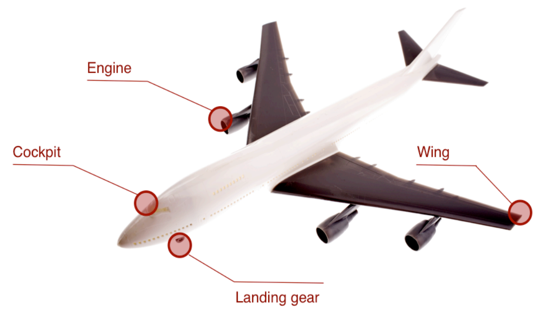

Fig. 1. Actual airplane model used for the test. The highlights show the 3D areas that can be recognized by the system when the plane is scanned under the webcam.

III. Technology

Several technologies and two main software components have been employed to implement the prototype:

- feature recognition
- visual search

The feature recognition system is designed to turn a digital image into the description of the set of its characterizing components. Corners, blobs, straight lines, T-junctions, curves and surfaces, matte colored areas: all of these can prove to be interesting, recognizable elements of a physical object. The current system is based on a weighted combination of all these possible features that proved to be effective in identifying the set of toys we experimented on.

Technically, the digital image coming from the webcam is pre-processed using a series of filters (saturation, color balancing, exposure, thresholding) and it is then fed through a three-stage process: first the object is isolated as much as possible through a simple background-removal algorithm; then an integral histogram of the distributions of color on the various area of the object is calculated; and lastly, a SURF (Speeded-up robust features) algorithm is applied.

The result of this first phase is composed of a set of metadata describing the identified features and color distributions and localizations on the object.

This information is used as input to the visual search software component. Here, the features and color distribution data are interpreted as vectors on n-dimensional spaces whose dimensions are the quantities involved (axes, for example, can be the localization or intensities of colors, or the spatial placement of the features). These vectors are superimposed on the corresponding ones obtained by pre-processing the physical object. In this pre-processing phase, the object (the toy) is captured along a series of three dimensional positions, in order to allow its recognition under multiple interesting perspective views. Vectors are then matched for proximity and scaling: when the vectors corresponding to the image captured by the webcam are sufficiently close to the set of vectors representing the toy, the matching occurs.

On a consumer Mac OSX laptop equipped with a hybrid Processing/Objective C environment, and using the OpenCV, P-SURF, and OpenFrameworks libraries, we have achieved a level of successful identifications of around 82-90% in different lighting conditions, at 10-12 frames per second on a 640x480 webcam capture.

IV. Conclusions

Toys++ aims at creating an open learning environment that could be easily manageable by educators, that will be free to define specific educational content, conveyed through any kind of personalized embodied agent.

References

[1] S. Papert, Mindstorms, Basic Books, 1991.

[2] E. Ackermann, “Piaget’s constructivism, Papert’s constructionism: What’s the difference,” Future of learning group publication, 2001.

[3] Digital Box, http://www.ohgizmo.com/2009/01/24/legos-digital-box-kiosks-feature-augmented-reality/, accessed 2010-01-03; Avatar i-Tags”, http://www.avataritag.com/, accessed 2010-01-03; Junaio, http://www.junaio.com/, accessed 2010-01-03.

[4] A. Nijholt, “Agents, believability and embodiment in advanced learning environments,” Proc. IEEE International Conf. on Advanced Learning Technologies, 2001, pp. 457–459.

. Dede, “Immersive Interfaces for Engagement and Learning,” Science (New York, N.Y.), vol. 323, 2009, pp.