This document is published in:

2011 IEEE Global Engineering Education Conference (EDUCON): Learning Environments and Ecosystems in Engineering Education (2011). IEEE, 1171-1176.
DOI: http://dx.doi.org/10.1109/EDUCON.2011.5773295

Towards Parallel Educational Worlds

Carlos Delgado Kloos, Carmen Fernández-Panadero, Mª Blanca Ibáñez, Mario Muñoz, Abelardo Pardo
Department of Telematic Engineering, Universidad Carlos III de Madrid
28911 Leganés (Madrid/Spain)
{cdk, mcfp, mbibanez, munozm, abel}@it.uc3m.es

Abstract— **Augmented Reality, 3D virtual worlds, etc.: the technology has evolved tremendously and so has its application to the field of education. Digital technologies have advanced to the point, where we are reproducing digitally more and more aspects of our life. We have parallel worlds: on the one hand the real world, and on the other virtual worlds, that can in fact be linked to the real one. They have different properties, but they can enrich and complement each other. In this paper, we explore the possibilities and challenges of these parallel worlds for educational uses.**

Keywords: e-learning, technology-enhanced learning, learning management systems, 3D virtual worlds, mirror worlds, augmented learning

I. Introduction

Technology applied to education has already been used for more aspects then just to overcome distance, as it used to be a long time ago. Three decades ago, the discussion was about tele-education or remote learning. But now, without forgetting about distance learning, we talk about blended learning, since technology can also enhance face-to-face education. After a class with the teacher in a conventional classroom, the student could go home and continue with learning activities through a learning management system and a networked computer. But now this technology-enhanced part of the learning has been experimenting interesting advances. We'll mention here two of them:

1. mobile, pervasive, ubiquitous access to learning any time and anywhere
2. 3D virtual environments, that emulate with increasing precision the face-to-face setting

In relation to the first point, the technology enhancement can happen anywhere thanks to the use of mobile devices. Several scenarios can be thought of. For instance, students could provide the teacher with up-to-date feedback about their understanding of the explanations in class (clickers mode) or look up complementary material. Mobile devices can also be used to aid learning in field trips. Through augmented reality techniques, mobile devices can overlay a layer of virtual objects or information on top of real world environments. Etc.

In relation to the second point, the technology for 3D multi-user virtual environments or metaverses is now ripe for a conventional computer to render 3D worlds where avatars representing learners and teachers can interact in a learning environment. The worlds can represent the real world (mirror worlds) or be completely invented (virtual synthetic worlds). Or they can be a mixture. They can even look out to the real world: one could have in the virtual world a panel showing the image that comes from a camera in real time.

The last examples (augmented reality and 3D metaverses with real time camera view) show that the two environments, the real and the virtual one, are not independent. Therefore learning could not only happen in the real or virtual world alone, it could happen in a mixed environment. This opens up interesting venues and new learning scenarios. Learning designs are not just bound to the real or the virtual world; they can take advantage of both worlds, in *parallel educational worlds*. This was already the case with blended learning, but now what is real and what virtual gets intermixed and blurred.

In this paper, we will study three different topics: learning objects, assessment, and learning activities. For each of these topics, we will review what they mean in traditional terms and what they can imply under the new idea of parallel educational worlds. In particular, we will study how the traditional description tools are applied to the new learning scenario that includes augmented reality and virtual worlds. We will do this in sections III, IV, and V, reporting about the work in progress and the challenges that lie ahead. Before that, in section II, we will the review the state of the art in linking the real world with a virtual one. We will conclude in section VI.

II. Linking the Real World to Virtual Worlds

The term "virtual world" can refer to a physically non-existing world, such as the one we experience in a dream. But in this paper when we talk about virtual worlds, we will always refer to digital worlds, as those generated by computers.

Education occurs at moments in the real world, and at moments in a virtual world. But the parameter "location" is not made explicit and the virtual image of a person is just reduced to a user login.

When talking about virtual worlds, we can distinguish clearly two elements in them: the avatars and the environment. The avatars represent the people (learners, teachers). The environment is everything else we can find. The existence of several avatars distinguishes 3D virtual worlds, also called MUVEs (multi-user virtual environments), from 3D simulations that just have one player. The possibility of having multiple users coexisting in the space in synchronization allows having collaboration, which is so important for learning.

How can we link objects (including persons/avatar) in the real world to a virtual world? There are a number of technologies that allow making this link. Here are a few:

- Geo-location
- Graphic tags: QR codes
- Electronic tags: RFID, NFC
- Recognition in the visible, IR, or ultrasound spectrum
- 2D or 3D projections
- Holography

Geo-location: The geographical coordinates (longitude, latitude, altitude) identify the place of a static object and therefore also to some extent the object itself. We can use these geographical coordinates in a digital world to refer to the real-world object, and so establish a link.

Text tags: When we cannot easily establish the geographical coordinates of a real-world object (for instance, because we are in the interior of a building and have not access to GPS information) or if the object does not have a permanent position, we can physically attach a text to an object. This text can be a URL or a SLURL (or some other 3D reference) or, even better, a URI (uniform resource identifier). The text acts as an interface between the digital world and the real-world object that has it attached.

Graphic tags: Instead of attaching a text directly, we can code it graphically. There are many formats to encrypt texts into graphics: linear bar codes, or two-dimensional ones, such as QR codes, Data Matrix codes, or EZcodes, to mention a few. There are many apps for handheld devices that can read these codes.

Electronic tags: A more complex tag that can be attached to an object is an electronic tag. Electronic tags, implemented using technologies such as RFID or a particular flavor of it compatible with mobile devices such as NFC, are capable of storing chunks of data readable and sometimes re-writable using radiofrequency waves. Electronic tags use globally unique identifiers that make them traceable and locatable providing the capability to create anchor points in the physical world that can be linked to virtual worlds.

Near Field Communication (NFC) technologies were developed in 2002 by Philips and Sony and adopted by ECMA International as a standard in December 2002 and by the International Organization for Standardization (ISO) and the International Electrotechnical Commission in 2003. The NFC Forum [1] was created by Philips and Sony together with Nokia in 2004. As of November 2010, the NFC forum had grown to over 130 members. The technology behind NFC operates using radio-frequency communications in the 13.56 MHz band, which is widely used for communication with Radio Frequency Identification (RFID) tags and smart cards. NFC has been made compatible with RFID based contactless communications. NFC-enabled mobile devices can emulate ISO 14443 smart cards and act as active devices in peer to peer (P2P) scenarios. As captured in [2], NFC defines a mechanism by which wireless mobile devices can communicate with peer devices in the immediate locality (up to 20 cm), rather than rely on the discovery mechanisms of popular short-range radio standards. NFC technologies provide a natural way to embed RFID reading and writing functionality inside a mobile device providing a user friendly way to access information chunks scattered in the user's physical environment without the need of external elements such as presented in [3].

An even more sophisticated way to identify objects is by some sort of *recognition*. This recognition can take place in different ranges of the electromagnetic spectrum. Camspace [4] uses the *visible* spectrum and can turn everyday products like cans or bottles into controllers replacing the mouse, keyboards, and joysticks. Some game consoles like Nintendo Wii [5] use the *infrared* spectrum to identify the user's position. And there is some research looking for new uses of this kind of devices like the Wiimote or other IR cameras [6]. Also many of the pens [7], whiteboards [8], and virtual keyboards [9] present on the market use IR technology to identify the position of a physical object.

Recently some work has also been exploring the spectrum of *ultrasound* using acoustic sensors that pick up sounds of low frequency that sound when touching a surface and correspond to an action in the system. This technique has been traditionally used in flat screens, but Sensitive Objects [10] (now belonging to Tyco Electronics [11]) has extended this concept to any surface using Anywhere Multitouch Platform™ [12]. Microsoft in collaboration with Carnegie Mellon University has applied the same concept to the human skin [13].

Most modern game consoles combine several of these techniques such as in the case of Sony Kinect™ for Xbox 360 [14], which combines IR technology with image and voice recognition. Most of these technologies have been used to transform daily life objects into tactile interfaces [15] in order to introduce real objects in the virtual world.

But there is also a huge amount of possibilities when we do the opposite: projecting virtual objects into the real world.

The 6th Sense project [16] is an example of sophisticated ways to use *2D projections* to obtain additional information about an object (e.g. a book, a wall), a person in real time, or even to project onto our body some device that we need like a calculator or a clock. There has been some research [6] to adjust the projection of surface that folds in 3-dimensional space.

The *projection of 3-dimensional* images is a more complex issue. There have been some initiatives to solve this problem: Stereoscopic systems that show two simple 2D images for the left and for the right, which are integrated by the brain. This technology requires special devices, typically glasses. Images are shown with slightly different perspective for each eye. There are different technologies to get this, like anaglyph images (different color layer for each eye), polarized glasses (different polarization for each eye), chroma depth codifies the depth of the objects modifying the way colors are perceived in each eye and LCD shutter glasses that block vision alternately in each eye while shows images with a slightly different perspective in the other one. Other technologies are based on perception imperfections like the Pulfrich effect (a dark lens over one eye that has the effect that when something moves from left to right, it'll look like it's moving back or forward).

All these early generation stereoscopic displays require the viewer to wear special glasses. To avoid the use of glasses there has been research in auto-stereoscopic devices. Some of these devices also use optical trickery at the display because they use a series of simple 2D images, creating 3D looking effects that also take advantage of eye imperfections. The problem using imperfections of human visual perception is that it can cause various side effects, like jumping images, limited field of view or headache.

Another common technology to provide a 3D-experience is the use of *headtrackers* combined with a flat screen that show different perspectives depending on the user's head position. This effect has problems with the number of simultaneous users looking the display. To avoid these effects, the Graphics Lab at the University of Southern California [17] creates 3D images using a high-speed video and projecting it onto a quickly spinning mirror. The mirror reflects a different and accurate image in different angles to each potential viewer. The system uses an algorithm to figure out the correct shading and occlusion for the image. But these technologies also have the problems derived of optical trickery.

The only technology that meets all the expectations of true three-dimensional images is *holography*. Holography does not uses the way in which image is perceived but also the way in which is produced. In this technique, the 3D light field scattered by an object is registered and then reconstructed. With this technique we can record real or virtual objects. There are lots of displays depending on the techniques used to record, register, and reproduce the image [18]. One of the most popular holographic displays is Holovizio™ [19]. This device consists of a monitor that generates all the light beams to make the 3D view of the displayed object visible in the whole field of view. To do that, it uses voxels instead of pixels. Each voxel can project multiple light beams –of different intensity and colors– in several directions, simultaneously. This means that anyone standing around the monitor will actually see an object from a different perspective, with no need for goggles or other stereoscopic tricks. Holographic images need not be static. There are results in the production of holographic video [20].

In most of these techniques, we see the 3D light field as coming from the object. But there is another technique to concentrate the light beams to place where the object would be rather than disperse them as if they came from it. The problem of this technique is that a special medium is needed to reproduce the image. One of the most important results of this technique has been obtained in Japan using a plasma-laser hologram device [21] that takes advantage of the plasma emission phenomenon near the focal point of focused laser light. By manipulating the laser's focal point, along the x, y and z axes, they can display real 3D images in mid-air.

With this review of the state of the art, we see that technology is quite mature to combine real objects with virtual ones, offering many interesting linking possibilities among worlds. Now let's turn our attention to learning objects.

III. Learning Objects

The concept of "learning object" is an interesting one. The reference to "object" in "learning object" (henceforth LO) could imply that LOs are physical entities that can be touched. IEEE defines an LO as "any entity, digital or non-digital, that may be used for learning, education or training" [22]. So an LO can be non-digital, but most of the times when we talk about LOs we refer to digital chunks of information, such as zip-compressed files with media resources as prescribed in SCORM.

The introduction of tangible objects with physical behaviour that can be touched, picked up, moved, and so one, evokes other senses for the user during the learning process. Using neuro-linguistic programming terminology [23][24], in addition to the auditory and visual representation systems, users are allowed to use the kinaesthetic representation system that takes into account physical sensations and associated emotions, and this should increment motivation and therefore improves learning.

Let's consider the following scenarios, all related to learning languages:

- 0: A real life performance of three people speaking
- 1: A written conversation among three people (read on a book)
- 2: A written conversation among three people (read on a tablet)
- 3: A cartoon animation with three characters speaking
- 4: A movie with three people speaking
- 5: A 3D virtual world with three avatars speaking

It is clear that all of them are variations of the same idea. There are some differences, though:

- 0 and 1 have a real context and 2 to 5 a virtual one
- In 0 and 5, the learners (or their avatars) can participate in the conversation, making the learning more effective. 0 requires physical presence, whereas 5 has the advantage of not requiring it.

With the introduction of 3D virtual worlds, the possibilities of interaction with digital objects are enormously enriched by new forms of interaction. Traditional LOs, such as SCORM packages, are document-oriented. These packages offer great versatility allowing access by multiple users and adapting their multimedia contents to their respective needs. But user behaviour is limited to traditional web interaction: upload files, introduce text, and manipulate hypertext and multimedia controls using the keyboard and mouse.

The LOs in the 6 cases above should be similar from the technological point of view. This is easy for 1 to 5, were the beholder has no digital representation and the content is text or video: just a different mime-type. But with 6 the *content* becomes *context*. If content could be easily isolated and packaged and offered to the users through an LMS, the situation becomes a bit more complicated. One has to identify what is common to several activities and users and what are the differences that have to be packaged as LOs to be presented at the right time for each learner. The packaging should make no problems, since all resources are typically XML files, but one has to identify what to factor out.

IV. Assessment

IMS QTI (Question and Test Interoperability) [25] is an interoperable assessment specification that has been used to evaluate knowledge in traditional Learning Management Systems. The assessment tests are composed of a set of questions with a pre-determined correct answer and possibly also a feedback and a grading scheme. Question types are multiple choice, multiple response, true/false, fill-in-the-blank, etc. These are all mechanisms adapted to "flat" interfaces.

When we move from flat web pages to 3D worlds, suddenly a wealth of rich possibilities appears [26][27]. Instead of *selection* actions out of a given set of possibilities, many possible actions in a 3D world can be used to represent the response from the user; for instance, moving a particular place, doing something with a particular object, interacting with a particular avatar, etc. Rather than asking whether a screw is inserted clockwise or anticlockwise (as text options) and let the user tick the right response, one can ask the user to perform the right action. Rather than asking the learner which of some predefined sequences to assemble a system represents the right one, one could ask the learner to assemble the system in the 3D environment. Not only the question and response part is much richer, also the feedback can be much more expressive. It can be shown how the screw enters clockwise into the wall, or how a system is assembled in the correct way.

Moreover, not only knowledge can be assessed, but also skills and competences [28][29]. The environment for assessment in a 3D world resembles much more the real world, and therefore the assessing power is much greater. One is not restricted to talking about actions; one can do the actions directly (virtually).

But let's move form a virtual 3D world to the real world. A gymkhana is the natural extension of formative assessment to the real world. Now, the same sequence of activities could be carried out in the physical world (say as a gymkhana, where the learner has to solve some skills-based challenges in the real-world), or in a virtual world (its virtual counterpart), or in a combination of the two (for instance, in a real-world with augmented reality information coming from mobile devices).

Similar assessments should be described with similar linguistic tools, independently of the environment they are carried out. Depending on the resources available at a particular moment in time and setting, one can enact them in one way or another (with more or less virtuality), without changing the initial specification. This flexible way of carrying out assessments can only be done, if the description is defined in a common way, independently of how the objects are instantiated (in the real world or a virtual one).

V. Learning Activities

What has been said for assessments in the preceding section, should be applicable for all kinds of learning activities, and therefore for the whole learning experience. Let's concentrate now on the possible pedagogies that can be used for a particular learning experience and more concretely the description of the learning designs.

Educational modeling languages (EMLs) are used to define the flow of learning activities to be carried out, as well as the roles played by the different stakeholders and how these build groups at different moments. Several EMLs have been proposed, although no best solution has been found so far. IMS LD (Learning Design) [30] is based on the metaphor of a theater performance. These are the main elements of IMS LD:

- A *play* represents the main (root) element and may contain a number of *acts*, which are run in sequence.
- Within an *act*, there are *role-parts*, which are run in parallel. This allows teachers and learners to perform different *activities* at the same time, or arrange the learners in groups.
- An *activity*, which may be either a *learning activity* or a *support activity*, has a number of *parts.*
- Each *part* can have *learning objectives, prerequisites,* and *meta-data* and makes reference to an *environment.*
- An *environment* contains the *LOs* and *learning services* to be used in an *activity.*

When considering virtual worlds, the metaphor of the theater performance and many of these elements become much more adequate, since the avatars play in fact the role of actors that have to perform certain tasks. It is even possible to distinguish the different roles by certain complements attached to the avatars, as e.g. is proposed in [31]. In this paper, the authors propose to enrich the avatars attire with complements such as moderator hats, jigsaw shirts (each jigsaw group with a different color), expert group jackets, etc.

With IMS LD and for an LMS, each activity occurs in a particular environment, which contains certain LOs and learning services. In the same way, a virtual platform supporting IMS LD should offer for each activity the right environment, implemented through a 3D scenario, with the correct LOs and learning services. The different successive environments can be mapped to different 3D scenarios or alternatively to the same with LOs and learning services appearing and disappearing according to the advance of the learning. A combination of these two alternatives is also possible. We are presently working on the programmed appearance and disappearance of LOs and learning services following IMS LD units of learning [32].

Now, a particular IMS LD play may be enacted either completely

- in the real world (LOs are tangible objects, environments are physical rooms with these objects, learners –in person– can meet in groups in separate parts of the room –same time, same place).
- in a classical LMS (LOs are digital, document-oriented resources, environments can be web pages that give access to these LOs, learners –represented by user logins– can meet in groups in separate digital spaces,

such as wiki pages or email threads –note that this meeting can be asynchronous).

- in a 3D virtual world (LOs are rich 3D virtual resources, environments are 3D scenarios, learners –represented by avatars– can meet in separate 3D zones).

It gets more interesting when we combine these possibilities. This combination can occur in several ways:

- One world at a time: The same play occurs at times in different worlds. An example of this is *blended learning*: in one act, the teacher explains something in the real world, and in the following act, the learners in groups write essays collaboratively about the explained concepts through an online tool, say a wiki.
- Learners in different worlds: Part of the learners follow the explanation of the teacher in a face-to-face meeting (real world), while another part, physically on a distant place, follow the explanation through digital means (be it through videoconference –remote learning– or a 3D virtual world).
- Combination of worlds: Learning activities are enriched by combining input from different worlds. An example could be a stroll through a city learning about history and seeing images of buildings and monuments as they looked in the past, by using augmented reality techniques.
- All the previous together: An example can be the gymkhana mentioned earlier, where some learners are in the real world and some in a 3D mirror world. The real world learners get augmented information coming the virtual world and vice versa.

VI. Conclusion

In this paper, we have explored how to unify the description of different settings from the same perspective. Many things that are apparently different become much similar under this common light.

It is clear that the real and the virtual worlds are useful to have side by side. Combinations of worlds can help in understanding relationships and therefore assist in learning [33]. It is more flexible to postpone the decision on how to implement each activity. Then, depending on various circumstances, such as resources available, devices used, etc. one would be able to enact them accordingly. In order to postpone this decision, one needs to be able to describe the designs, assessment, etc. in common formats. Most of the existing formats serve very well as a basis for the new scenarios and technologies. It also becomes clear that the "*location"*-parameter (in the widest sense) has to be made explicit in the description language and instantiated as late as possible.

Acknowledgment

The eMadrid Excellence Network is being funded by the Madrid Regional Government (Comunidad de Madrid) with grant No. S2009/TIC-165. We wish to acknowledge stimulating discussions with our partners in the context of the network. Partial support has also been received from the Learn3 project (TIN2008-05163) and the SOLITE project (CYTED 508AC0341).

References

(Latest access for all web sites: December 2010)

[1] www.nfc-forum.org/home

[2] R. Want: *An introduction to RFID technology*, IEEE Pervasive Computing, Vol. 5, No. 1. (2006)

[3] J. Riekki, T. Salminen, I. Alakarppa: *Requesting pervasive services by touching RFID tags*, IEEE Pervasive Computing, Vol. 5, No. 2. (2006)

[4] www.camspace.com

[5] Nintendo Wii, wii.com

[6] Johnny Chung Lee: *Projector-based location discovery and tracking*. Ph.D. Thesis. Human Computer Interaction Institute. School of Computer Science. Carnegie Mellon University, May 2008. johnnylee.net/projects/thesis/

[7] www.iogear.com/product/GPEN200NW6

[8] www.labwe.com/touchscreen.html

[9] I.tech virtual keyboard, www.vkb-support.com/learn_more.php

[10] www.sensitive-object.com/spip.php

[11] www.tycoelectronics.com

[12] www.sensitiveobject.fr/pdf/ EN_PR_Sensitive_Object_Anywhere_MultiTouch.pdf

[13] C. Harrison, D. Tan, D. Morris: *Skinput: Appropriating the body as an input surface*. Proc. CHI '10: 28th Annual SIGCHI Conf. Human Factors in Computing Systems, Atlanta, GA, 10-15 April 2010, ACM, New York, NY

[14] www.xbox.com/en-US/Kinect

[15] Alan Crevoisier, Cédric Bornand: *Transforming daily life objects into tactile interfaces*, Proc. Euro-SSC 2008: 3rd European Conf. on Smart Sensing and Context

[16] www.pranavmistry.com/projects/sixthsense

[17] Andrew Jones, Ian McDowall, Hideshi Yamada, Mark Bolas, Paul Debevec: *Rendering for an interactive 360 light field display,* ACM SIGGRAPH 2007, gl.ict.usc.edu/Research/3DDisplay/

[18] www.zebraimaging.com

[19] www.holografika.com

[20] alumni.media.mit.edu/~lucente/holo/holovideo.html

[21] *Japanese device uses laser plasma to display 3D images in the air*, PhysOrg.com, 27 Feb 2006

[22] IEEE Learning Technology Standards Committee: *Draft standard for learning object metadata*. IEEE Standard 1484.12.1, New York: Institute of Electrical and Electronics Engineers, p. 45. (2002)

[23] Joseph O'Connor, John Seymour: *Introducing NLP*, The Aquarian Press 1990

[24] Robert Dilts, Todd Epstein: *Dynamic learning*, Meta Publications, Capitola, CA 1995

[25] IMS Global Consortium: *Question & Test Interoperability Specification v2.0/v2.1*. www.imsglobal.org/question/

[26] D. Morillo Arroyo, P. Santos Rodríguez, D. Pérez Calle, C. Delgado Kloos, M.B. Ibáñez Espiga, D. Hernández-Leo: *Assessment in 3D virtual worlds: QTI in Wonderland*. Congreso Iberoamericano de Informátia Educativa, IE2010, Santiago de Chile, Chile, 1-3 Dec. 2010

[27] Mª Blanca Ibáñez, José J. García, Sergio Galán, David Maroto, Diego Morillo, Carlos Delgado Kloos: *Multi-user 3D virtual environment for*

Spanish learning: A Wonderland experience, ICALT 2010, 5-7 July 2010, Sousse, Tunisia, IEEE Press

[28] Mª Blanca Ibáñez, Raquel M. Crespo, Carlos Delgado Kloos: *Assessment of knowledge and competencies in 3D virtual worlds: A proposal*, IFIP World Computer Congress 2010, Key Competences in the Knowledge Society, Brisbane, Australia, 21-23 Sept. 2010

[29] Mª Blanca Ibáñez, Raquel M. Crespo, David Aguado, Carlos Delgado Kloos: *Enhancing acquisition and distribution of knowledge in professional environments using 3D virtual worlds*, MATEL workshop, ECTEL2010, Barcelona, Spain, Sept. 2010, Springer

[30] IMS Global Consortium: *Learning Design Specification v1.0.* www.imsglobal.org/learningdesign/, 2003

[31] Thrasyvoulos Tsiatsos, Theodouli Terzidou: *Supporting collaborative learning processes in CVEs by augmenting student avatars, with nonverbal communication features*, ICALT 2010, Tunisia (2010)

[32] David Maroto, Carlos Delgado Kloos, Derick Leony, Mª Blanca Ibáñez, José J. García: *Orchestrating learning activities in 3D virtual worlds*, unpublished, 2011

[33] Guillaume Zufferey, Patrick Jermann, Son Do-Lenh, Pierre Dillenbourg: *Using augmentations as bridges from concrete to abstract representations*, in: Proc. 23rd British HCI Group Annual Conference on HCI 2009: *Celebrating People and Technology*, pp. 130-139, Swinton, 2009. British Computer Society