The Potential of Kinect in Education

Hui-Mei Justina Hsu, Member IACSIT

Abstract—This paper explores the potential of Kinect as interactive technology and discusses how it can facilitate and enhance teaching and learning. Kinect is examined in terms of its affordances of technical interactivity, which is an important aspect of pedagogical interactivity. As it utilizes gesture-based technology, Kinect can support kinesthetic pedagogical practices to benefit learners with strong bodily-kinesthetic intelligence. Though it has facilities to fabricate meaningful classroom interactions, Kinect technology can not stand alone in the classroom setting but needs to be integrated with a computer, projector and compatible software. As far as a teaching tool is concerned, due to the multiple interaction types it supports, Kinect has the potential to enhance classroom interactions, to increase classroom participation, to improve teachers’ ability to present and manipulate multimedia and multimodal materials, and to create opportunities for interaction and discussion. As a learning tool, Kinect has the affordances to create enjoyable, interesting interactions types, to boost student motivation, and to promote learning via its multimedia and multi-sensory capacity. In addition, students can utilize the bodily information gathered by Kinect with software programs to create highly interactive multimedia works. However, the implementation of Kinect in the classroom has technical constraints such as large classroom space, lack of easy-to-use development tools, and long calibration time and pedagogical constraints such as the difficulties in shifting to kinesthetic pedagogical practices and limited understanding of its effect.

Index Terms—Bodily-kinesthetic intelligence, gesture-based computing, Kinect, educational technology.

I. INTRODUCTION

Kinect is described as a revolution in the making as it provides a brand new type of interaction with computers [1]. It is a motion sensing input device by Microsoft for the Xbox 360 video game console, which can capture, track and decipher body movements, gestures and voice. The auditory and visual information serves as commands to interact with digital contents presented in games or software programs. In other words, users are not bound by keyboards, mice or joysticks and thus have intuitive and virtual experiences with digital contents. The kinesthetic and gesture-based interaction enabled by Kinect definitely is the dream application that computers are envisioned to support.

The original context of Kinect’s application is in combination with Microsoft Xbox 360 for the purpose of enhancing entertainment experiences. Thanks to hackers’ success and the release of open source drivers, Kinect’s potential is greatly enhanced [2]. Kinect’s release aroused hackers’ attention, and assumptions were made that radical new applications would be possible if Kinect could be made to communicate with computers. Soon after the release of Kinect in Europe, it was successfully hacked by Héctor Martín, who had produced a Linux driver that allowed the use of both the RGB camera and depth sensitivity functions of the device [3]. It was not until Prime Sense’s release of the open source drivers along with motion tracking middleware called NITE that the prevalent application of Kinect started to take off. In addition, Microsoft released a non-commercial Kinect software development kit (SDK) for Windows in July 2011. The applications of Kinect in various fields become easy to implement.

The appearance of Kinect certainly draws the attention of educators and encourages the evaluation of its feasibility in education. For example, Kissco [4] expresses his excitement by predicting that Kinect will become a focal classroom technology in the next few years. Similarly in the 2010 and 2011 Horizon Report [5], [6], gesture-based technology such as the Microsoft Kinect system is believed to be ready for adoption in the educational setting in four to five years. This paper, therefore, would like to explore Kinect in terms of its potential in facilitating and enhancing teaching and learning. The analysis intends to provide an examination of its affordances and emerging pedagogical opportunities.

II. AFFORDANCES

In spite of the strong optimism for Kinect, the literature of educational technology (e.g., [7]) has cautioned us to take a critical perspective for fear that technological promise should be more than its reality. The benefits of integrating technology into the classroom include meeting the needs of visual learners, more interactively teaching whole-class lessons, and better engaging students [8]. The introduction of ICT resources to schools, however, seems to have relatively little impact on the ways that teachers teach [9]. Consequently, due to the strong optimism for the capabilities of digital tools in transforming students learning outcomes, there is a tendency to overestimate the learning outcomes that a given medium or technology will automatically guarantee [10].

In order to evaluate educational technology properly, the technology (the tool within the setting) should be discussed in relation to the pedagogy within a subject area (the practice in the setting) and the subject domain and its culture (the ecology of the setting) [10]. Such a cultural aspect echoes with Bowers’ [11] comments on the amplifications and reductions occurred in the process of technology integration. In other words, the technological affordances allowed for certain technologies would shape and facilitate certain
Interactivity is the most perceived advantage of ICT in terms of its support for teaching. Interactivity is defined as the ability to respond contingently to learner’s actions [13]. The term ‘interactivity’, therefore, can be used to describe technical interactivity as technology serves as an interface between the user and the material, and pedagogical interactivity, which is itself a teaching strategy [14]. In order to enhance the interactivity of classroom activities, technology needs to provide interactive affordances, which can be implemented to support interactive teaching strategies and to create pedagogical opportunities.

Interactivity with ICT can be classified into five categories: no interactivity with ICT, authoritative interactivity, dialectic interactivity, dialogic interactivity and synergistic interactivity [13]. No interactivity with ICT describes the use of technology as merely presenting digital contents. In the case of authoritative interactivity, students are controlled by predetermined responses with fixed procedures. From dialectic interactivity to synergistic interactivity, students are allowed more freedom to formulate their desired interactions with contents and technologies. Dialectic interactivity constructs an environment where multiple resources are pulled together to support student exploration and hypothesis-testing.

In addition to auditory and visual modalities, current technologies start to react to pressure and motion of fingers and to take gesture-based inputs. The availability of multi-sensory presentation increases interactivity as the human actions are better afforded by features of ICT. In the case of interactive whiteboards (IWBs), they serve as a fortified version of chalkboards, allowing users to manipulate physically with text and image. Furthermore, when writing on IWBs, users can easily write in different colors with finger touch, erase contents, move and resize objects and utilize the library of images provided by the IWBs software. Moreover, an IWB is actually an extension of a computer. Users can control the classroom computer via the connected IWB. IWBs enable students and teachers to communicate more intuitively with tactical force and body movement.

IV. BODILY-KEINASTHETIC INTELLIGENCE AND GESTURE-BASED COMPUTING

The interactivity facilitated by Kinect has to do with gesture and body movement. Its educational foundation can be traced back to the idea about the three primary learning modalities: auditory, visual and kinesthetic. Kinesthetic learners constitute 15% of student population [14], and they learn better when they touch or are physically involved in what they are studying. Therefore, teaching methods need to be properly modified in order to cater the needs of these learners.

In addition, the Theory of Multiple Intelligences articulated by Howard Gardner [16] provides another educational rationale for the accommodation of Kinect in the classroom. The MI theory is an approach to understanding human intelligences with many educational implications. Different from the previous notion of intelligence as a single capacity, the MI theory assumes a pluralistic cognitive universe [17]. In MI theory, human intelligences are classified into eight types, and each of them represents a relatively independent form of information processing [18]. Among them, bodily-kinesthetic intelligence, which describes one’s abilities to control body movements and to handle skilfully, has a lot to do with the educational potentials provided by gesture-based computing. As students come to class with different intelligences, kinesthetic approaches to teaching are used to facilitate teaching subjects such as English (e.g. [15], [19]). With the assistance of body movements and gestures, meaning making seems to be grappled more easily by students who originally have learning difficulties.

Gesture-based computing can be viewed as an innovative educational development in alignment with bodily-kinesthetic intelligence. It involves devices controlled by natural movements of the finger, hand, arm, and body [20]. Due to the fact that current technologies begin to incorporate touch to computer-assisted learning environments, it becomes crucial to understand how touch and hands-on experiences enhance understanding of concepts and what aspect of learning outcomes is salient when incorporating the tactile and kinesthetic features of ICT.

Haptic information is claimed to have a positive impact on learning as shown in recent research literature from the field of psychology [21]. For example, an experiment studies how children age of five learn phonetic rules [22]. In addition to visual and auditory clues, the researchers ask the children to trace words with an index finger while pronouncing the word and looking at it. The result suggests that such a multi-sensory approach proves to be effective in establishing
the connection between the visual and auditory representations. Moreover, the addition of haptic feedback is believed to create a more ‘real’ experience and is strongly recommended in educational simulations. It is thought that hands-on experience would allow students to have a deeper understanding of concepts.

A recent study is conducted at Nottingham Trent University by Rachael Folds, a PhD researcher from the University’s School of Education, experimenting on the effect of the Nintendo Wii and Xbox Kinect in helping college students with learning difficulties. The participants are college students aged from 16 to 24 with disabilities ranging from Down’s Syndrome to autism spectrum disorders. Folds asks one group of participants to play tennis using a Wii tennis game and another group to play bowling using a Kinect game. The results show that with five week of computer game training two groups of participants improve significantly comparing the pre- and post-test scores in real activities [23].

The application and development of gesture-based computing continues to thrive in the field of training and education. 2011 Horizon Report [6], for example, lists several on-going projects, such as DepthJS and 3Gear System at the MIT. DepthJS is a Kinect project allowing users to interact with web browsers by gestures; 3Gear System makes use of Lyvra gloves and computer cameras to trace the movement of hand. With the release of open source drivers and non-commercial Kinect software development kit, it becomes easy to take advantage of the facilities of Kinect in the field of education.

One problematic aspect about the applications of haptics and gesture-based computing in education is that researchers are not yet clear about how haptics improves learning. The advantage of haptics is believed to do with kinesthetic memory especially when students are asked to perform complex, three-dimensional motor skills [21]. But it is not clear how learners process kinesthetic information and use it to improve understanding. The related research studies and developments are still in a primitive stage. Intensive interdisciplinary studies and collaborations are required in order to exploit the full potential of kinesthetic and gesture-based technologies. Without solid research findings, such applications may not favoured or last long.

V. AFFORDANCES OF KINECT IN THE CLASSROOM

Though it has the potential to facilitate natural interactions, Kinect needs to be situated in combination with software and other hardware in order to fabricate meaningful classroom interactions. Compared with IWBs (price ranges from about US $800 to US $2,500), Kinect is relatively cheap as it costs around US $149. If the classroom is equipped with a projector and a computer, Kinect can be regarded as an inexpensive add-on.

So far the applications for Kinect in education are still in a developmental stage. Recently, a German company, Evoluce, utilizes Kinect depth sensor and develops a software program, WinandI, to enable gesture control over Windows 7 desktop, browsers, Microsoft Office and other applications. With WinandI, teachers are flexible in terms of their positions to computers because Kinect can actively track their movements and gestures within four meters to the screen. If WinandI or similar software programs can be made easily accessible, most things that can be done with IWBs can essentially be done with Kinect.

However, similar to the case of IWBs, the facilities of Kinect are largely dependent upon the software used. So far there have been few handy software programs for teachers to create Kinect-enabled contents. If Kinect comes with software for teachers to design the control over computers, it would surely become a powerful interactive educational technology. Based upon its facilities and preliminary applications, the following analysis intends to depict Kinect’s potential in two broad categories: a tool to enhance teaching and a tool to support learning

A. A Tool for Teaching

Kinect and IWBs share some similarities as both intend to enhance the level of interactivity of the current technology-enabled classroom. The potential benefits of IWBs for teaching derive from six major themes: flexibility and versatility, multimedia/multimodal presentation, efficiency, supporting planning and the development of resources, modeling ICT skills, and interactivity and participation in lessons [24]. In terms of interactivity, students seem eager to come up to the board and enjoy interacting physically with IWBs. As the kinesthetic and gesture-based aspect is the major characteristics of Kinect, it can be predicted that with the accommodation of Kinect there will be more pedagogical activities tailored to the needs of kinesthetic learners in the classroom. For example, students are encouraged to interact virtually with objects or to act out scenarios depicted in readings. The gesture-based affordances enable teachers to design instruction to involve students’ kinesthetic intelligences and to meet the needs of kinesthetic learners. In order to exploit the potential of Kinect, however, school curricula need to make a shift to kinesthetic pedagogical practices.

As to the impact on student participation, the practice of body as a controller will have students immersed more in the simulated learning environment. Even in the IWB-enabled activities, students feel like a magician and thus have a strong willingness of participation [24]. Connected to computers and projectors, Kinect can provide natural and intuitive interaction patterns via body, gesture and voice control over digital contents. Just as the current applications in gaming, Kinect’s verisimilitude will create the sense of reality and have students find the similarities between interactions in the simulated environment and those in the reality.

The incorporation of Kinect will enhance teachers’ ability to present and manipulate multimedia and multimodal materials. In the case of IWB technology, teachers are empowered by the capabilities of highlighting, annotating, dragging, dropping and rotating text and image. In combination with visual and auditory information, teachers are able to create multi-sensory interactive activities with Kinect, because it can easily integrate body and gesture movement into classroom activities [24]. Consequently, teachers are allowed to utilize a good combination of students’ multiple intelligences.
Another potential benefit is the increase of opportunities for interaction and discussion. In order to exploit Kinect’s technical affordances, classroom activities should be designed to use the information gathered by Kinect. It follows that the pedagogical strategies should encourage student participation by body movements, gesture and voice. In the context of IWBs, ESL teachers ask students to walk to the whiteboard and place words in the correct column by dragging the cursor [8]. Such a pedagogical strategy is reported to be useful as it allows students time to engage in the mental representation of the target words and provides a space for discussion. Kinect can present similar or even better pedagogical opportunities and can be viewed as a viable alternative for IWBs. For example, instructions can be designed to have students work on a physical procedure, which allows students time to use body movement to test their hypothesis.

In terms of a tool for teaching, Kinect has the following four characteristics. First, Kinect is a flexible teaching tool. Teachers can use or design contents with interaction of body movements, gesture and voice without using keyboards or mice; in other words, they are more flexible about their locations and do not need to stand close to the classroom computer.

Second, Kinect can accommodate multiple users. So far the Kinect games in Microsoft Xbox 360 can have two people playing at the same time. Therefore, students and teachers can have a fare share of control over interactions. A Kinect-enabled classroom can support whole-class instruction, group work and teacher-student one-to-one interaction.

Third, it is a versatile tool. As it collects three-dimensional information, Kinect can support various teaching activities such as dance and martial arts. Special instructional design can be implemented to reinforce the connection between teaching contents and student physical responses. For example, when students see a clue, they need to act out in order to proceed. If designed properly, the interactive contents can enable students’ vicarious participation in physical activities.

Fourth, Kinect engages students. Compared with those enabled by IWBs, the interactions enabled by Kinect support multiple physical engagement patterns, involve more time on task from students, and imply better utilization of multiple intelligences.

B. A Tool for Learning

As a tool to support learning, the affordances of Kinect can be analyzed in three major aspects. First, it is a stimulating tool. One major advantage claimed in the IWB literature concerning the benefits of learning is motivation and affect [24]. Kinect can be integrated into simulated environments and greatly enhance their verisimilitude. If lesson plans and interactions are carefully designed, the Kinect-enabled classroom should have the affordances to create enjoyable and interesting interaction types to boost student motivation.

Second, another feature of Kinect to promote learning is its multimedia and multi-sensory capacity. Kinect facilitates kinesthetic interactions and is able to coordinate with visual and auditory information to reinforce student learning. As they are able to interact with contents physically, students can make use of kinesthetic memory to assist recall. Not only does more information become available, the ideas and concepts become more tangible and easier to grasp.

Third, Kinect can be used with software programs to enhance its role as a learning tool. The idea of a learning tool aligns with constructivism, which emphasizes building external, sharable artifacts and personal relationship with knowledge in the process of learning [25], [26]. Educational software is designed to promote the construction of personal representations of knowledge. Due to the fact that Kinect can gather information from users, students can add creativity to their multimedia works by feeding the information into the programs. In this way, Kinect can extend the varieties of interaction types supported by the software programs and bring new features to the multimedia works created by students. Two creativity tools, Scratch and Mikumikudance, are demonstrated in regard to their application of Kinect.

Scratch (http://scratch.mit.edu/) is an educational programming language developed at MIT Media Lab. It is a learning tool as well as a cognitive tool for kids to explore different subjects of interest. Scratch can connect to sensors such as Picoboard and LEGO WeDO for different types of interaction with the real world.

Thanks to the effort of Stephen Howell, Scratch is able to utilize the capacity of Kinect. Howell combines the accessibility of Scratch with Kinect to allow children to physically interact with a software program without having to touch the screen, the keyboard or a mouse. The body information containing the x-, y- and z-axis values captured by Kinect can be fed into Scratch via OpenNI2Scratch program. In total, information concerning 15 body parts, including the head, shoulders and hands, is collected by Kinect sensors (as shown in Figure 1).

With the x-, y- and z-axis values of body parts, students can program Scratch applications to allow interesting interactions with users. For example, KinectSkeleton (shown in Figure 2) is a game created by Stephen Howell, in which Scratch, the main character, maps the positions of the player’s skeleton. The program consists of loops that constantly check the position of the player’s body parts and reflect the changes to the position of the colored circles.

With Kinect sensors, children are provided with rich bodily information that they can use to create their stories or games with fun and good interactions. In a workshop with elementary school students, the researcher finds that students
can easily make games such as hitting and throwing balls and gesture magic shows to utilize the capacity of Kinect.

![Fig. 2. Kinect selection program](image)

Mikumikudance is another creativity tool utilizing the capacity of Kinect. Mikumikudance, commonly abbreviated to MMD, is a free animation program for novice animators to create 3D animations. This program has been popular among young people for its easy and quick implementation. With the advent of Kinect, a plug-in is made and expanded by Mr. Mogg to get the motion of neck, wrist, upper body, lower body, and ankle in MMD Kinect motion capture. Instead of manually manipulating joints, animators can make use of Kinect plug-in to do a quick 3D animation by acting out. Figure 3 is a screen capture of a 3D model mapping the movements of the user (the red man). The addition of Kinect to MMD lowers the technical threshold for younger students to express their creativity via 3D animation.

![Fig. 3. MMD program with Kinect plug-in](image)

VI. CONSTRAINTS

Implementation of Kinect has technical constraints as well as pedagogical constraints.

A. Technical Constraints

There are four major technical constraints. First, the implementation of Kinect requires enough empty space to the screen, which may be difficult to find in a regular classroom. In addition, the position of Kinect may become problematic as it needs to be situated in front of the players. In other words, a stand may need to be placed between students and the whiteboard.

Second, calibration takes time, which may result in the waste of instruction time. To begin with, it may take up to a few minutes for Kinect to track a new user. Furthermore, sometimes situations arise for re-calibration when students walk out of the range that Kinect can reach or when calibration is not done correctly. Moreover, when the number of players increases, the procedure of calibration may disrupt the class.

Third, there are still very few software programs and teaching materials available that utilize the kinesthetic potential of Kinect. Despite the release of open source drivers and the non-commercial SDK, still very few content developers incorporate the Kinect’s technological features.

Fourth, in order to make good use of a multitude of information gathered by Kinect, including the information generated from multi-user scenarios, some easy to use programs should be available for teacher to make good use of the bodily information and to generate educational contents.

B. Pedagogical Constraints

Despite of the fact that the application of Kinect has its educational foundation in multiple intelligences, visual and auditory modalities have long been established as the major ways of learning at school. Consequently, it may be difficult to persuade teachers to modify their instructional design in a kinesthetic fashion. In addition, the current experiments and developments with Kinect are still in a primitive stage. Researchers are not clear about how kinesthetic or multi-sensory interactions can enhance learning. Even if it is motivating for students to engage in a kinesthetic activity, a lesson plan without tangible objectives may not result in apparent learning outcomes. Last but not the least, the integration of Kinect into learning requires adequate training for teachers in order to use it to achieve teaching efficiency. Teachers will need to be trained to set up Kinect and manipulate it features and to construct lesson plans with kinesthetic features. However, if support is not available or the teaching efficiency is not enhanced, teachers’ enthusiasm may quickly wanes.

VII. SUMMARY AND IMPLICATIONS

The above analysis suggests that Kinect has great potential to enhance classroom interactions and to ignite student creativity. Kinect technology, however, can not stand alone in the classroom setting but needs to be integrated with a computer, projector and software. Customized software to facilitate classroom interactions and to create Kinect-enabled contents seems to be missing in the picture of current technology integration. Therefore, the evaluation of Kinect as the focal classroom technology largely depends on the future development of Kinect software. Software needs to incorporate the design of interactive pedagogy in order to exploit its technical interactivity.

Judging from the preliminary applications, Kinect is capable of being a tool to enhance teaching and a tool to support learning. This paper draws the connections between Kinect and IWBs and believes that Kinect can provide better interaction types and fulfill most of the facilities offered by IWBs. One advantage that is absent in IWB literature is that Kinect has the facility to gather information, which can later be fed, into student creativity tools. Its application in Scratch and MMD serves to diversify student representations of
knowledge. Future studies are needed to address student creativity enabled by Kinect. In addition, due to the fact that researchers are not yet clear about how the application of haptics and gesture-based computing improves learning, Kinect needs strong empirical evidences to support its legitimacy as an adequate classroom technology.

REFERENCES


Hui-mei Justina Hsu obtained her Ph.D. from University of Illinois at Urbana-Champaign in 2006, majoring in curriculum and instruction. She is currently working as assistant professor at the department of learning and digital technology in Fo Guang University in Taiwan. Her research interests are technology integration, cultural aspect of educational computing and children programming.

Dr. Hsu has been a member of IACSIT and IEEE since 2010. She has attended various academic conferences in the field of educational technology and serves as presenter as well as keynote speaker.