Gesture-Based Interaction for Seamless Coordination of Presentation Aides in Lecture Streaming

Andrew K. Lui, Vanessa S.C. Ng, and Chun-Hong Chan

School of Science and Technology, The Open University of Hong Kong,
Good Shepherd Street, Ho Man Tin, Kowloon, Hong Kong SAR, China
alui@ouhk.edu.hk

Abstract. This paper describes a novel lecture streaming system that has integrated gesture based interaction with a flexible coordination of presentation tools. The system aims to facilitate instructors giving a seamless lecture delivery that can focus more on engaging students. With simple to learn gestures, the instructors can conveniently call upon presentation tools such as presentation slides, student response system, and electronic whiteboard. The system can be used in an office with minimum hardware and no technical staff support, allowing the arrangement of impromptu lectures at a distance to promptly share any inspiration thought about an emerging influential event. The paper also reports some positive evaluation results.

Keywords: Lecture streaming, natural interaction, gesture, kinect, direct manipulation, distance learning.

1 Introduction

Over the Internet streaming of lectures provides opportunities for remote students to enjoy some unique benefits of attending a lecture in person. Delivering lectures with inspirational and critical thoughts are known to promote motivation (McKeachie 2002) and to facilitate development of problem solving skills, especially in computing and mathematics (Boster et al. 2006). Current learning content delivery technology have come a long way from traditional distance education, and it is now beginning to see live or delayed lecture streaming has replaced its videotape or digital versatile disc (DVD) counterparts. Several research indicated that the performance of students exposed to lecture streaming were either better than or on par with those attending face-to-face lectures (Flower & Sawa 2006; Boster et al. 2006; Buhagiar & Potter 2010).

Lecture streaming has been increasingly adopted among tertiary education institutions in the world. It costs significantly less in production and distribution (Shephard 2003). A computer device with a connection to the Internet is the only requirement for joining a highly scalable learning network that allows a wide distribution of cost associated with a distance-learning programme. There are also readily available software that allows the recording and reviewing of lecture video

J. Lam et al. (Eds.): ICT 2013, CCIS 407, pp. 108–119, 2013.

clips. These benefits, including low cost, high-perceived value, scalable operation, and broadened admission market, have convinced senior university management that lecture streaming has a promising future.

A particularly attractive usage scenario of lecture streaming is impromptu lecture at a distance. An instructor can promptly share any inspiration thought about an emerging influential event. For example, a teacher in liberal studies can speak about an ongoing political crisis or a professor in geology can discuss about a current volcanic activity. Lecture streaming can also be integrated with online discussion forums. For discussion topics that are difficult to explain with text, giving a short presentation can be a more effective solution. Lecture streaming can be coupled with video recording so that the discussion video can be made available for delayed viewing as well as real time broadcasting.

Given the increasing maturity of computing technologies, instructors should be able to begin lecture streaming from their own office in the near future. It is important that an instructor could manage the whole lecture streaming process individually. The absence of technical support staff would save cost. The instructor would enjoy the freedom of starting a lecture at any time and the convenience of not having to go to a lecture theater or a production studio.

1.1 A Gesture Controlled Lecture Streaming System

The purpose of this paper is to describe the design and development of a prototype lecture streaming system for individual instructors' use in their office. The prototype system is the first attempt of a project that aims to investigate effective and efficient ways to deliver lectures through video streaming.

A key design concept of the system is to use *gestures* as the means of controlling various operations of lecture streaming. Gestures are movement of body parts, and in the context of communications the hands and upper limbs are usually involved. They form a crucial part of human expression and they can carry meanings beyond the capability of speech and words (Roll 2001; McNeill 1992). It is therefore not surprising that gestures play an important role in teaching and learning (Roll 2001; Lanir et al. 2008; Arnold 2012). Modern inspirational lectures use gestures cleverly to maintain student engagement (Andersen et al. 1979), coordinating interactions (Ford 1999), and facilitating concept construction (Lanir et al. 2008). Gesture-based input technology should allow a more seamless integration into lecture delivery.

With the lecture streaming system proposed, instructors can manage the whole lecture streaming session on their own. They can control various system operations with hand gestures instead of remote control devices or computer mouse or keyboard. They can call upon call up various presentation aides such as presentation slides and student response system, and arrange these items in different layout configurations in the video stream.

Our principal interest with the prototype system focuses on finding an effective design of gesture input interface. The desired objectives include identification of suitable hardware and general infrastructure support, and derivation of principles of the interaction design that does not hinder typical delivery of lectures.

1.2 Related Work

Natural Interaction and Gestures. Natural interaction describes a concept that human users can interact with a computing device based on human senses only. Current peripheral input devices such as keyboard or mouse would become obsolete if a computer user can interact naturally with a computer similar to that with a human being. Natural interaction would require significantly less learning time and effort (Wachs et al. 2011). Many smartphones have already incorporated gesture based natural interaction, and even toddlers can quickly know how to use their fingers to flip through photographs and to draw pictures. Microsoft Kinect, a motion sensing device for gesture based games, offers a low-cost natural interaction hardware support for a number of innovative applications, such as sign-language recognition (Lang et al. 2012), guided imagery based relaxation system (Lui et al. 2012), and stroke rehabilitation (Pogrzeba et al. 2012).

Gesture Based Teaching and Learning Systems. Interactive whiteboard systems were the first teaching and learning technology that incorporated natural interaction. Basically, such systems use motion sensing devices to replace the sensory on conventional physical whiteboard surface. (Bosetti et al. 2012) described a whiteboard system based on a Wiimote that could detect drawing on a conventional project screen with a pen shape special device. Ronchetti & Avancini (2011) eliminated the need for any hand-held device and developed a Kinect based whiteboard that could be drawn with a fingertip. Because of noise and other environmental factors, these systems generally suffer from precision problems, and extensive calibration process would be required for achieving a reasonable usability level. Other interesting work in this area includes the one described in (Smorkalov et al. 2013), in which the captured gestures of a lecturer are sent into a 3D virtual world for controlling presentation slides and an interactive whiteboard. Wang et al. (2004) developed an innovative automated lecture video editing system. The region of interest of a presentation slide is inferred from the gestures capture from the lecturer and accordingly a close-up view of the region is shown in the edited video.

Access to Tools and Services with Gestures. There are several approaches of how a gesture based user interface provides access to various tools and services:

- Mapping specific gestures or postures to functions. Wang et al. (2004), for example, designed a circling gesture for specifying the region of interest. There is a straightforward one-to-one mapping between a gesture and a function. If there are many functions then remembering all the gestures may be quite challenging.
- Direct manipulation. Bosetti et al. (2012) and Ronchetti & Avancini (2011) both designed a hand-controlled virtual mouse for directly manipulating tools in their interactive whiteboard system. A problem with direct manipulation is that any gesture or involuntarily movement is interpreted as input. Both systems required additional means of invoking or switching off a function with a handheld pen or a virtual mouse click.

– Menu system. A menu system can provide an effective access to a potentially large set of functions. The invoking of a function is becoming a step-by-step process, and at each step a few choices are presented for selection. A few specific gestures are now mapped to choices and do not pose a problem in remember all the gestures.

2 Method

The research objectives addressed in this paper include the following:

1. Design an effective means of accessing functions in a lecture streaming system with gestures.
2. Design and implement a prototype lecture streaming system for proof-of-concept and evaluation.

The project work began with an analysis of functional and non-functional requirements. Then an initial design of a gesture based user interface was drawn up. The prototype system was designed and implemented to allow a preliminary evaluation of the gesture based user interface. The preliminary evaluation provided useful data for refining the design of the user interface and the final prototype system.

2.1 System Design Overview

Figure 1 below shows an overview of the lecture streaming system. The software part of the system is running on a typical PC computer connected to a Microsoft Kinect, and the PC together with the Kinect form the hardware part of the system.

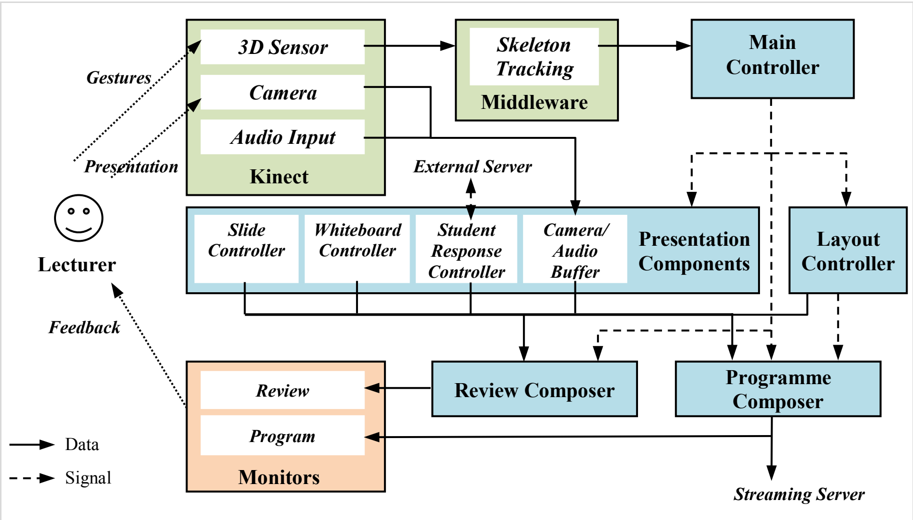


Fig. 1. Overview of the system design of the lecture streaming system

2.2 Hardware

Microsoft Kinect is the key hardware component, providing motion sensing data as well as audio and video frame capture. Therefore it is doubled up as not just a gesture input device and also a video camera. The Red-Green-Blue (RGB) camera offers VGA resolution (640 x 480) at 30 frames per second, which is more than enough for lecture streaming.

The principal Kinect feature relevant to this project is skeleton tracking. With the help of a middleware such as OpenNI[1] or Microsoft Kinect SDK[2], various joints and body features of one to four users may be tracked in the 3D space. To get into skeleton tracking, the middleware would first require the user to be posed in some particular way for skeleton calibration.

Kinect is a suitable choice for use in an office. The effective range of both the camera and skeleton tracking is 1.2 meters to 3.5 meters. The viewing angle of the cameras spans 43 degrees vertically and 57 degrees horizontally. It can be adjusted with a tilt motor.

2.3 Software

A specially designed program aims to analyze the camera and skeleton data streams from Kinect, and to produce the final programme for lecture streaming based on the desired layout configuration that specifies how various presentation aides are composed on the screen. The program has the following major components:

1. Presentation component manager. A presentation component has a visual appearance that can be part of the composition of the lecture streaming programme. Examples of presentation components include:
   - Camera (showing the lecturer)
   - Presentation slides
   - Student response system
   - Interactive whiteboard
2. Layout controller. It has built-in several standard layouts for arranging the presentation components such as split-screen, picture-in-picture, etc.
3. Review composer. It provides visual feedback when users are interacting with various tools and services using gestures.
4. Programme composer. It receives instructions from the layout controller and the main controller and generates video frames for lecture streaming.
5. Main controller. It maintains the states of the whole system and sends signals to other components according to the current state and gesture input.

---

[1] http://www.openni.com

[2] http://www.microsoft.com/en-us/kinectforwindows/

2.4 Gesture Based Interaction Design

Gesture based interaction is adopted for accessing various tools and services in the lecture streaming system. There are two types of tools and services: system tools and presentation tools. The two system tools include the *layout controller* described above and the *configuration manager*. The latter would allow users to control the streaming process and to change the settings of system configurations. Provision of access to the presentation tools would also be necessary. For example, moving to the next presentation slides or drawing on the electronic whiteboard would be essential in a lecture streaming session.

The gesture based interaction design for this system has a two-level structure as shown in Figure 2 below.

At the top level, there are typically four to six system tools and presentation tools that require access. Both the *direct gesture mapping* and the *menu system* approaches have been considered in the project. In the *direct gesture mapping* approach, each of the six tools is assigned a unique gesture for invocation, and the same gesture is used to close the tool (Figure 2a). On the other hand, the *menu system* approach specifies one gesture (ie. the *invoke gesture*) for switching from one tool to the next, and specifies another gesture (ie. the *exit gesture*) to exit from the tool (Figure 2b). The *menu system* approach is adopted in the project based on the negative feedback on the direct gesture mapping approach collected from the preliminary user evaluation (refer to the next major section).

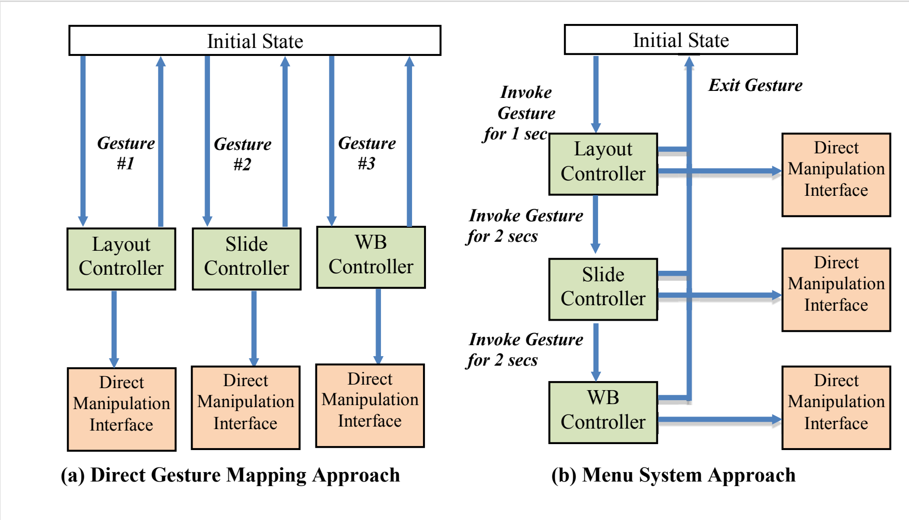


Fig. 2. The two-level gesture based interaction design

At the bottom level, each tool provides its specific user interface. The direct manipulation approach is chosen in the design because it has a universal applicability similar to a computer mouse pointer. The hand position is mapped to the pointer and a *press* action takes place when the pointer stays in a region for one second. The *review monitor* provides the necessary feedback to users.

The top level and the bottom level gesture tracking should not interfere with each other because one uses the left hand and the other uses the right hand. Figure 3 below shows in the menu system approach how the graphical user interfaces changes with holding on the *invoke gesture* for a few seconds.

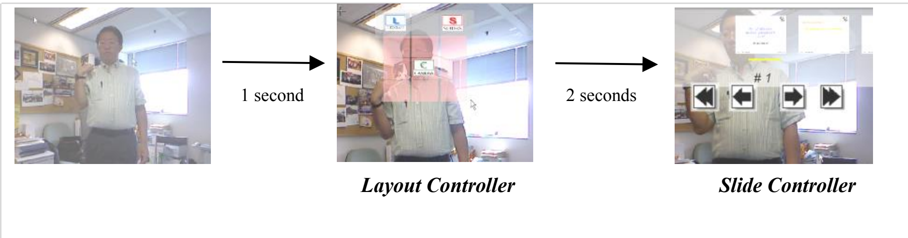


Fig. 3. An illustration of how to access the Slide Controller by holding onto the Invoke Gesture

On the actual definition of the *invoke* gesture, the principle to follow is to minimize the interference to the lecture delivery as much as possible. Table 1 lists the definitions of the *invoke* and the *exit* gestures. The particular poses are selected because they are not typically seen in teaching context. Figure 4 below shows photographs of the two gestures.

Table 1. Definition of the invoke and the exit gestures in the gesture menu system

Gesture names	Remarks
Invoke	Left hand placed near the left shoulder with the left elbow tugged in to the waist.
Exit	Left hand placed near the right shoulder with the left elbow touching the torso.

Fig. 4. Illustrations of the *invoke* and the *exit* gestures

2.5 Prototype Design and Implementation

A prototype lecture streaming system has been designed and implemented. The program part is written in Java and it has integrated the following libraries and components:

- Java Media Framework (JMF) for video processing and streaming.
- Apache POI, the Java API for Microsoft Documents, for the import of powerpoint files.
- Likeclass, an in-house student response system, which is external to the lecture streaming system and connected by HTTP and JSON.

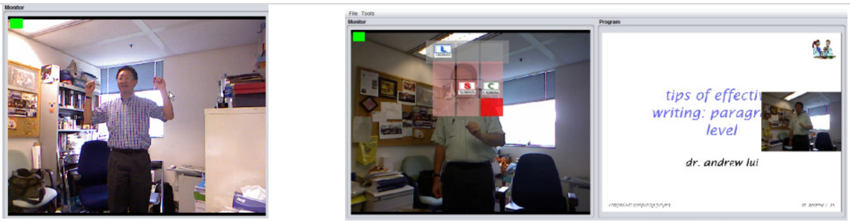
A dual-core 3.0 GHz personal computer with 2 GB of RAM was found to be sufficient for a smooth running of the system.

3 Results

The purpose of this section is to illustrate, first, that the proposed design of the gesture based lecture streaming system is feasible in the implementation and functional aspects; and second, that the gesture based interaction design is proper.

3.1 Operational Aspects of the System

Figure 5b shows the layout of the graphical user interface of the prototype system. The left and right windows are the *review monitor* and the *program monitor* respectively. The *program monitor* shows the final video stream as seen by viewers.



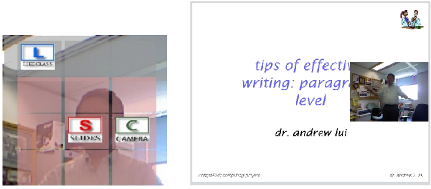
(a) Calibration Pose ***(b) The Layout of the GUI of the program***

Fig. 5. (a) The calibration pose for skeleton tracking; (b) The layout of the program's GUI with the Program Monitor showing a picture-in-picture layout configuration

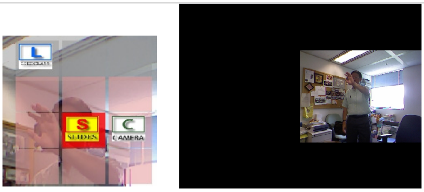
To begin using gesture based interaction, the instructor must perform a calibration pose (figure 5a). A green indicator appears if skeleton tracking is active. After that, the screen layout, system configuration, and presentation tools can be controlled by gestures.

Figure 6 illustrates how the layout of presentation components can be directly manipulated with a hand. A red square will follow the position of the hand to assist the navigation. Staying on an icon which represents a presentation component, for more than one second will pick up the component. The icon will turn yellow and will follow the hand position. After the icon is moved to the desired grid position, it can be dropped by pulling back the hand.

Figure 7 shows how direct manipulation is applied to presentation slide control. There are four buttons for moving the slide show forward or backward in two different speeds.

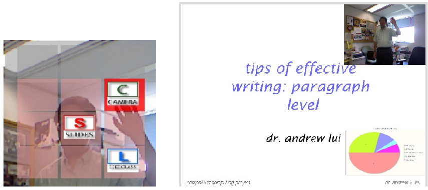


(a) The component (Slide) occupying the centre grid position takes up the whole screen. Other components are placed accordingly



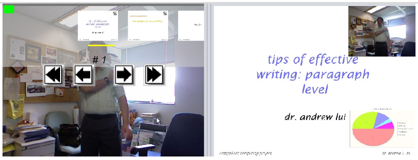
(b) Placing the hand over the Slide component will select the component (turned yellow)

(c) Placing the Slide component to the left grid position has turned the layout the left-right split

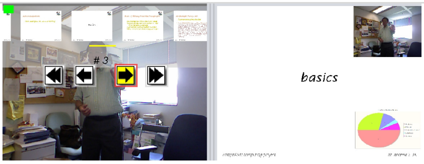


(d) Student Response component and the camera are now placed to the corners

Fig. 6. Illustration of using the Layout Controller to change the layout of presentation components



(a) The four buttons for slide control are shown on the screen



(b) Pressing and holding the Forward button has moved to the slide after the next

Fig. 7. Illustration of using the Slide Controller

3.2 Evaluation

Five instructors were invited to participate in system evaluation. All five were instructors of computing courses but only two of them had a little experience with gesture based games.

At the beginning of the evaluation, a two-minute briefing was provided to explain the purpose of the prototype system. It was then followed by a two-minute explanation of how to gain access to various tools with gestures, and how to use direct manipulation to interact with the *layout controller* and the *slide controller*. A two-page tutorial notes with visual illustrations of the *invoke* and the *exit* gestures was also provided.

Study #1: Direct gesture mapping as opposed to gesture menu

The objective is to evaluate the direct gesture mapping approach, which is described in Table 2. The participants were asked to call upon the layout controller and then asked to exit the layout controller. Then they were asked to do the same with the slide controller.

Table 2. Definition of the gestures for calling upon and exiting the layout and the slide controller in the direct gesture mapping approach

Gestures	Tools mapped
Left hand placed near the left shoulder with the left elbow tugged in to the waist.	Layout controller
Left hand placed near the right shoulder with the left elbow touching the torso.	Slide controller

Three out of the five participants showed hesitation, confusion, or error with remembering the correct mapping. One participant voiced out the difficulty in remembering the right gestures and suggested a method similar to the current menu system approach. There is little semantic relation between the two gestures and the tools they are mapped to.

Study #2: Competency in controlling the layout

The participants were asked to carry out a sequence of tasks with the layout controller. Everyone completed all the tasks within a minute. Table 3 shows the average time taken for every task. It also compares the performance of participants experienced in gesture control to those without experience.

Table 3. Mean time taken to complete tasks with the layout controller

Tasks	Inexperienced (s)	Experienced (s)	Overall (s)
Pickup a component	3.7	3.0	3.4
Pickup and replace a component	7.7	5.0	6.6
Move Slide component to right position	3.3	4.5	3.8
Move Likeclass component to bottom left	6.3	3.0	5.0
Make the camera inactive	6.7	4.5	5.8
Arrange 2 components in a left-right split	8.0	9.0	8.4

Most participants were able to smoothly carry out the tasks, with the occasional hiccups, of which the most frequent one was the replace a component to a location. The performance difference between the experienced and the inexperienced was found to be insignificant. In fact the inexperienced participants showed a remarkable skill improvement within the one minute period.

Study #3: Competency in controlling the slides

The participants were asked to carry out a sequence of tasks with the slide controller. Table 4 lists their performance.

Table 4. Mean time taken to complete tasks with the slide controller

Tasks	Inexperienced (s)	Experienced (s)	Overall (s)
Move one slide forward	5.3	4.0	4.8
Move to slide #6	5.0	3.5	4.4
Move one slide backward	2.7	2.0	2.4
Move to slide #2	2.3	2.0	2.2

Only one participant showed an initial problem with controlling the slide with the hand. The performance was found to be excellent, perhaps because the participants already went through the one minute practice with the layout controller.

4 Conclusion

Zarraonandia (2012) pointed out that lectures have been evolving with technological advancements in the past few hundred years. As a new medium of instruction delivery, lecture streaming would necessitate a new lecturing method for better engage remote students and facilitate active learning. The proposed lecture streaming system introduced a novel concept of using gestures to coordinate various presentation tools and it should allow a more seamless delivery of lectures without the need of onsite technical support.

This paper described a preliminary study into the feasibility of such gesture based lecture streaming system. The evaluation results were found to support that such system is natural to interact with and easy to learn. More experiments are required to find out the perception of such a delivery medium from the viewers' perspective.

References

1. Andersen, J., Andersen, P., Jensen, A.: The measurement of nonverbal immediacy. Journal of Applied Communication Research 7, 153–180 (1979)
2. Arnold, L.: Dialogic embodied action: Using gesture to organize sequence and participation in instructional interaction. Research on Language & Social Interaction 45(3), 269–296 (2012)
3. Bosetti, M., Pilolli, P., Ruffoni, M., Ronchetti, M.: Interactive whiteboards based on the WiiMote: Validation on the field. In: Proceedings of 14th International Conference on Interactive Collaborative Learning (ICL), pp. 269–273. IEEE (2011)
4. Boster, F.J., Meyer, G.S., Roberto, A.J., Inge, C., Strom, R.E.: Some effects of video streaming on educational achievement. Communication Education 55, 46–62 (2006)

5. Buhagiar, T., Potter, R.: To stream or not to stream in a quantitative business course. Journal of Instructional Pedagogies 3, 1–6 (2010)
6. Flower, E., Sawa, S.: Synchronous streaming video from the classroom to the remote student's desktop: Report of a distributed education project. In: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, vol. 2005(1), pp. 654–665 (2005)
7. Ford, C.E.: Collaborative construction of task activity: Coordinating multiple resources in a high school physics lab. Research on Language and Social Interaction 32, 369–408 (1999)
8. Lang, S., Block, M., Rojas, R.: Sign language recognition using kinect. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2012, Part I. LNCS, vol. 7267, pp. 394–402. Springer, Heidelberg (2012)
9. Lanir, J., Booth, K.S., Findlater, L.: Observing presenters' use of visual aids to inform the design of classroom presentation software. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 695–704. ACM (2008)
10. Lui, A.K.F., Wong, K.F., Ng, S.C., Law, K.H.: Development of a mental stress relaxation tool based on guided imagery and Microsoft Kinect. In: Proceedings Sixth International Conference on Distributed Smart Cameras (ICDSC), pp. 1–6. IEEE (October 2012)
11. McKeachie, W.J.: McKeachie's teaching tips: Strategies, research, and theory for college and university teachers, 11th edn. Houghton Mifflin, Boston (2002)
12. McNeill, D.: Hand and mind: What gestures reveal about thought. University of Chicago Press, Chicago (1992)
13. Pogrzeba, L., Wacker, M., Jung, B.: Potentials of a low-cost motion analysis system for exergames in rehabilitation and sports medicine. In: Göbel, S., Müller, W., Urban, B., Wiemeyer, J. (eds.) Edutainment 2012/GameDays 2012. LNCS, vol. 7516, pp. 125–133. Springer, Heidelberg (2012)
14. Prince, M.: Does active learning work? A review of the research. Journal of Engineering Education 93(3), 223–231 (2004)
15. Ronchetti, M., Avancini, M.: Using Kinect to emulate an Interactive Whiteboard. MS Thesis in Computer Science, University of Trento (2011), `http://latemar.science.unitn.it/segue_userFiles/LITSA/Avancini_Mattia_138793_Thesis.pdf` (retrieved)
16. Roth, W.M.: Gestures: Their role in teaching and learning. Review of Educational Research 71(3), 365–392 (2001)
17. Shephard, K.: Questioning, promoting and evaluating the use of streaming video to support student learning. British Journal of Educational Technology 34(3), 295–308 (2003)
18. Smorkalov, A., Fominykh, M., Prasolova-Førland, E.: Virtualizing real-life lectures with vAcademia and Kinect. In: Presented at the 2nd Workshop on Off-The-Shelf Virtual Reality at the International Conference on Virtual Reality (VR), Orlando, FL, USA, March 16 (2013)
19. Wachs, J.P., Kölsch, M., Stern, H., Edan, Y.: Vision-based hand-gesture applications. Communications of the ACM 54(2), 60–71 (2011)
20. Wang, F., Ngo, C.W., Pong, T.C.: Gesture tracking and recognition for lecture video editing. In: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 3, pp. 934–937. IEEE (2004)
21. Zarraonandia, T., Aedo, I., Díaz, P.: Envisioning the transformative role of IT in lectures. Interaction Design and Architecture(s) Journal - IxD&A (15), 7–17 (2012)