A Framework for MOOC Content Generation

Kanmanus Ongvisatepaiboon and Jonathan H. Chan
School of Information Technology, King Mongkut’s University of Technology Thonburi, Thailand
jonathan@sit.kmutt.ac.th

Abstract – Massive Open Online Course (MOOC) is becoming increasingly popular. There are different forms of MOOC and many online courses created by various institutions around the world. The traditional methods of generating high quality MOOC content tend to be time-consuming and are not easy for the lecturers to engage in natural interaction with the content. This paper presents a framework and proposes the use of the Chroma-Keying technique that would enable the lecturers to see the outcome in real-time and the use of simple hand gestures to control changing of slides.

Keywords: MOOC; Chroma Key; hand gesture; content generation

1. INTRODUCTION

The traditional pedagogical model for learning involves students going to the classroom to listen to the lecturer; this is termed “Teacher-Centered Education” [1]. With this model, lecturers typically assign students some assignments to test the students’ understanding of the course content. However, the model of learning system started to change some twenty years ago [2] with the advent of personal computers and improved network infrastructure. Students can now learn many things from the websites via online videos, e-Book, etc. There is no need for the lecturers to cover the theory as much in class anymore, so they change their role to be more like an advisor and/or facilitator. This is termed a flipped classroom [3].

Then Massive Open Online Course (MOOC) was created recently as another evolution of the learning method [4]. It provided videos, materials, problem sets, and forums that is very suitable for distance education.

Generally, for generating MOOC content, lecturers have to record their own face in a video and capture the screen separately. Then, they will put one video to be a layer above another one by using video editors. There is a limitation for this method that the lecturer cannot see the real outcome while they teach. As such, MOOC content is typically produced in a studio with additional post-editing required.

In this work, we propose a new method that is more convenient for lecturers to generate good quality MOOC content based on regular lectures that are recorded for e-learning purpose. This method will merge two videos by using the Chroma-Keying technique that will remove the background, similar to the weatherman’s forecast on television. Moreover, lecturers will be allowed to use simple hand gestures for controlling the presentation in form of a slide deck. A general framework to produce high quality MOOC will also be presented.

1.1. Video Merging

Generally, lecturers need to merge their own face video and screen capturing video after they finish their class. It means that they cannot engage in good interaction with the presentation (e.g. pointing to the text/image) because they cannot see the outcome in real-time. However, we have seen many newscasters stand in front of the scene and use their hand to point to a weather map. In fact, they stand in front of an empty scene with only one color, which is normally either green or blue because they differ most distinctly in hue from most human skin colors. This technique is called 'Chroma-Keying’ [5].

There are many applications that have a Chroma-Key effect. Some applications can show the live preview that can help lecturer to make a good interaction with the presentation (e.g. Chroma Key Live, VeescopeLive, and so on). However, if you want to make a high quality video with less noise, you should merge them carefully, after finish a lecture class, by using a professional video editor like Adobe After Effect, Adobe Premiere Pro, or iMovie.

Therefore, in this paper, we propose a framework that will use a live preview application and also record the lecturer’s face video and capture the screen separately. Using live preview application will help lecturer to see what the outcome will be. This can help them to have a good interaction. Then, they will merge their own face video and screen capturing video again for a high-quality finish.
1.2. Presentation Control

For controlling the presentation, normally people need to press a button on the keyboard or use a mouse-click. This requires the lecturer to be in proximity of the computer. As such, there have been much interest that tries to use hand gestures for controlling the slides [6][7]. One such app that allowed users to use hand gestures to control the presentation is “Flutter” (http://www.flutterapp.com). It supported both PowerPoint for Windows and Keynote for Mac. However, Flutter has since been acquired by Google and is no longer available as a standalone product.

1.3. Other Requirements

For our approach to generate a high quality MOOC content, camera and microphone are required. It can be a built-in device or USB device. We recommend to use a high-definition USB camera for high quality purpose.

2. FRAMEWORK

For generating high quality MOOC content, there are three main things that should be of concern: Video, Audio, and Presentation. The general outline of the proposed framework is shown in Fig. 1.

![Fig. 1. Framework for generation of high quality MOOC content](image)

2.1. Video

Video plays a very important part in MOOC. If the picture of the content is not clear enough, it may be difficult for the students to understand what the lecturer wants to convey [8].

For high quality, we should focus on the video camera device. According to Aventura Technologies, there are several factors we should be concerned with (i.e. bitrate, resolution, codec, data quality, and hardware) [9]. Generally, we are using a normal webcam or USB camera to record the video. At present, there are many high definition webcams and they are not overly expensive. From toptenreviews.com, they have been comparing different webcams annually. For the reviews on 2017 models, the top-rated one is Logitech HD Pro C920 for $58 and the cheapest model that can stream HD at 30 frames per second (fps) is Gear Head for only $18.96 (http://www.toptenreviews.com/computers/ peripherals/best-webcams/). That is, high-definition videos can be obtained affordably.

Most of HD webcams have maximum resolution around 1920x1080 (Full HD) currently. However, HD resolution of 1280x720 is likely sufficient for high quality MOOC content, depending on the screen size [10].

In this work, we propose a Chroma Keying method that will cut the background out and replace it with another video/image file. First, the scene’s background color should be different than the presenter’s outfit. The colors green and blue are commonly used because they differ most distinctly in hue from most human skin colors. This is the usual practice in studio production as well. About the software, configuration should be carefully modified. The quality of final result depends on the configuration of Chroma Keying effect.

2.2. Audio

Audio is another important part because students need to be able to hear the lecturers clearly. If the sound track is noisy, students may be annoyed and won’t understand what the lecturers want to describe [11].

For high quality audio, it depends on the microphone device of choice. There are several types of microphone. Basically, one can use a built-in microphone from a video camera or laptop. However, the overall sound quality is very low, so the sound is not very good. An alternative is using a USB microphone. The top-rated model from toptenreviews.com costs $249 and can rival studio sound.

A wireless microphone with a mini clip is another interesting choice. Since there is no need to carry the microphone, so the lecturer can make smooth movements with both hands.

2.3. Presentation

For our approach, lecturers should stand in front of an empty scene that may be far away from the computer. So, it would be awkward if they need to walk to click a mouse/keyboard. However, they can use a remote control, which is a better choice, but we think this is not a very natural movement. Therefore, we propose to use hand gesture recognition to make it more natural [12]. That is, lecturers can control their presentation by using their hands. Basically, we are focusing on some basic gestures for four functions (i.e. play/pause, move backward, move forward).
The accuracy of hand gesture recognition depends on the camera resolution and the algorithm [13]. The high definition video will make hand gesture recognition more accurate.

As for the algorithm, there are three main approaches for gesture recognition, which are 3D model-based algorithms, skeletal-based algorithms, and appearance-based models [14].

There are several approaches for recognizing hand gesture. In this paper, we will discuss two technologies: skeletal tracking with Microsoft Kinect and webcam-based gesture recognition.

Using Microsoft Kinect can detect the joint movement of human body [15][16]. This is a good choice for gesture recognition. Fortunately, there is a library “Kinect PowerPoint Control”, developed by Joshua Blake [7]. By using this library, we can control the PowerPoint presentation by using basic gestures.

Another method is using a general webcam for detecting hand gesture [17]. Since, webcam can be found easily and it is not too expensive. So, we recommend this approach.

Moreover, presenter’s skill is another important factor for generating a high quality MOOC content [18]. However, this is a kind of communication art, so we will not discuss this matter here.

3. METHODOLOGY

We separate the methodology section into three parts: Video Live Preview, Video Merging, and Hand Gesture Recognition.

3.1 Video Live Preview

In the traditional classroom, lecturers can write on the blackboard or show something via the projector. They can use their hand to point to the texts or images. For general online videos, they use a screen capturing method to show the presentation. They normally use a cursor to point what they want to make the audiences to focus on.

Since our approach is using the Chroma Keying technique to cut the background of lecturer’s face video out and replace it with a presentation. The lecturers need to see the live preview to guide them to make more natural interaction and movement. A TV screen may be mounted on the ceiling in front of the lecturer for this purpose.

Since our approach is using the Chroma Keying technique to cut the background of lecturer’s face video out and replace it with a presentation. The lecturers need to see the live preview to guide them to make more natural interaction and movement. A TV screen may be mounted on the ceiling in front of the lecturer for this purpose.

We use the ‘Chroma Key Live’ application (see Fig. 2) [19] for showing the live preview of lecturer’s face video merged with presentation. We choose this application because it is a simple one. We don’t need advanced features of professional software like Veescope, Wirecast, BoinxTV, or VirtualEyez.

This part is just for seeing the live preview. It will not record the result into a video file, since we want the high quality content. However, it provides a recording function, using a Syphon Recorder. So recording is possible using this approach if high quality is not an important concern.

Fig. 2. Sample screenshot of Chroma Key Live

3.2 Video Merging

While lecturers teach, they need to record their own face video and the presentation screen into separate files. Then, after finishing the class, they need to use a video editor application to merge them together by using a Chroma Keying effect. In this paper, we use ‘Adobe After Effect’ to merge them together.

The following are some briefly steps to merge two video files by using a Chroma Keying effect.

a.) First, import the lecturer’s face video with one-color background, which should be green or blue.

b.) Then, select the Chroma Keying effect by clicking in the menu Effect > Keying > Keylight (1.2)

c.) Next, change the screen color to be a background color.

d.) Can optimize by modify the Clip Black and Clip White function in the Screen Matte part. This will provide a high quality outcome.

e.) Finally, import the presentation screen to be another layer and put it below the lecturer’s face video layer.

3.3 Hand Gesture Recognition

Generally, lecturers are using an input device (e.g. keyboard, mouse, remote control, etc.) to control the presentation. However, hand gesture is another method for controlling the presentation. Since, there is no need to carry any device in their hand, so they can make a natural movement with the presentation.

There are many researches focused on hand gesture recognition [20][21][22]. There are some simple gestures with no movement: “Stop”, “Thumbs Up”, “Thumbs Down”, and “Point” [23].
In our framework, we recommend the ‘Flutter’ application, a gesture recognition startup acquired by Google in 2013. This application is supported in Mac and Windows. The Flutter’s main objective is to recognize hand gestures to control the applications (e.g. iTunes, YouTube, PowerPoint, Keynote, etc.) (see Figure 3).

Fig. 3. Flow chart of how Flutter app works

4. SUMMARY AND FUTURE WORK

The main objective of this paper is to propose a new method for assisting in the generation of high quality MOOC content. In our framework, we use a Chroma-Keying effect that will cut out the background of lecturer’s face video and replace it with a presentation video. Moreover, we propose to use a hand gesture recognition application to make it more convenient for handling the presentation. Therefore, lecturers can engage in more natural interaction during recording.

However, we think there are many features that can be added into this kind of work, e.g. complex hand gesture recognition. Also, we plan to obtain feedback of the system from lecturers as well as from students on the MOOC content.

References


http://www.zachpoff.com/software/chroma-key-live/


