Face Detection Technique as Interactive Audio/Video Controller for a Mother-Tongue-Based Instructional Material

Excel Philip B. Guidang, Christopher John R. Llanda, Thelma D. Palaoag
Lagangilang, Abra, Philippines, Lagangilang, Abra, Philippines, Gov. Pack Rd. Baguio City, Philippines
rieann12@gmail.com, john_llanda@yahoo.com, tpalaoag@gmail.com

Abstract. Face Detection Technique as a strategy in controlling a multimedia instructional material was implemented in this study. Specifically, it achieved the following objectives: 1) developed a face detection application that controls an embedded mother-tongue-based instructional material for face-recognition configuration using Python; 2) determined the perceptions of the students using the Mutt Susan’s student app review rubric. The study concludes that face detection technique is effective in controlling an electronic instructional material. It can be used to change the method of interaction of the student with an instructional material. 90% of the students perceived the application to be a great app and 10% rated the application to be good.

1. Introduction
Face Detection is the activity of automatically detecting and locating human faces in visual media or within an image, and returns face bounding boxes. It is the first-step in face recognition systems, with the aim of extracting the face region from the background [1]. Moreover, face detection algorithms targets on the discovery of frontal human faces. Face detection can be done in several ways such as scale-friendly deep convolutional networks, robust real-time face detection, neural network-based face detection, fast cascades, tracking-learning detection, subcortical face processing, component-based face detection, Eigenfaces, and Viola-Jones algorithm [2]. Among all face detection techniques, Viola-Jones algorithm if efficient and robust for detecting frontal faces [3] which was considered in this study. The process of detecting facial features can be utilized and applied in an e-learning environment.

E-learning refers to the teaching and learning process with the use of computerize medium to facilitate classroom based instructions. With the help of modern electronic devices nowadays ranging from small to large scale computers, teachers take advantage these technology to develop instructional material to be integrated in classroom instructions[4].

Modern instructions comes with blended learning with the integration of computer vision applications to interact students’ learning[4]. Face detection can be used in instructional materials especially on multimedia content topics to promote concentration and gives focus to learners. To further enhance the learning of the student, a multi-lingual capability of the multimedia was also considered in this study.

In Abra, where people have multiple dialects, tribe dialect also known as mother tongue language is used by people as the first language followed by Iloco. It is the commonly used dialect of people in Abra to communicate with other tribes. Filipino and English language is also used as a medium of instruction in the classroom. The problem is, students find lessons in programming difficult to understand when it
is delivered in English. So, language barrier is obvious. The need to develop an instructional material in mother-tongue is needed to eliminate language barrier.

Department of Education (DepEd) Memorandum No. 32, s. 2012, K-12 program obligates the use of Mother Tongue- Based Multilingual Education (MTB-MLE) in the delivery of lessons. The Abra State Institute of Sciences and Technology as partner of Dep-ED in the delivery of the K-12 Senior High School program offers TechVoc-ICT track on computer programming and Hardware servicing strand, finds the need to develop an instructional material adhering to the memorandum as it experienced language barrier. This study proves the possibility of replacing the method of interaction from keyboard and mouse inputs which are commonly used in e-learning by a face detection Face Detection Technique. The core idea of the application is to play the embedded instructional material and when it detects a face, suspend the same when it cannot, and resumes again when it re-detects a face. This study similar from that of Ugurlu, Y., (2012) in his study Human Interactive E-learning Systems wherein Head Posture Images was used as an input.

1.1 Objectives of the Study
The researchers present how face detection technique can be used in an Interactive Audio/Video Controller in an instructional material. Specifically, it achieved the following objectives:
1. Developed a face detection application that controls an embedded mother-tongue-based instructional material for face-recognition configuration using Python;
2. Determined the perceptions of the students and faculty members using the Mutt Susan’s student app review rubric.

2. Methodology
2.1 Research Design
The researchers used both descriptive and applied methods of research in the study. Descriptive research was used to narrate how the face detection application with embedded mother-tongue-based instructional material was developed and the perceptions of the students regarding the application. The study used Rapid Application Development (RAD) model to produce the Face Detection Technique as audio/video controller for an Instructional Material.

| Input                  | Process                        | Output                                      |
|------------------------|--------------------------------|---------------------------------------------|
| 1. Mother-tongue-based instructional material | Analyze and quick design Build Demonstrate Refine Test Implement | Face Detection Technique as audio/video controller for an Instructional Material |
| 2. Student perceptions      |                                | Feedback loop                               |

Figure 1. Paradigm of the Study

The paradigm in Figure 1 shows the interaction of the Input-Process-Output (IPO) variables. The input variables were mother-tongue-based instructional material and Student perceptions. The Face Detection Technique as audio/video controller for an Instructional Material was processed by the following procedures: a) analyze and quick design; b) build; c) demonstrate; d) refine; e) test; and f) implementation.

The output is the The Face Detection Technique as audio/video controller for an Instructional Material. The feedback loop of the study was based from the results of the demonstration being tested and evaluated by the users or respondents which provided room for improvement.
2.2. Sources of Data

Table 1. Distribution of Respondents

| Position          | N |
|-------------------|---|
| BSIT 4 Students   | 26|
| Faculty Members   | 4 |
| **Total**         | **30**|

Legend: N – total population

Table 1 shows the distribution of the respondents. Twenty six (26) BSIT students and four BSIT faculty members of the Abra State Institute of Sciences and Technology served as the respondents of the study. The respondents were asked to evaluate the face detection application with embedded mother-tongue-based instructional material.

2.3. Data Instrumentation

For objective 1, the application was developed using the Rapid Application Development model. It has six phases, namely, a) analysis and quick design; b) build; c) demonstrate; d) refine, e) testing; and f) implementation.

![Figure 2. Rapid application Model](image)

2.3.1. Analyze and quick design

Existing delivery method and platform of the Instructional material (IM) was examined. Moreover, viewing width of the devices used to access the School website was also considered.

The playing of the IM depends solely on the frontal face detection. The IM stops when frontal-face detection is not present and resume immediately soon after face detection.

2.3.2. Build

The mother-tongue-based IM was recorded using CamStudio and the face detection was developed using tracking.js framework. Later on, the IM was embedded in the face tracking framework.

2.3.3. Demonstrate

The application was shown to the some of the students while the application was being created to generate comments and suggestions for the improvement of the prototype.

2.3.4. Refine

As comments and suggestions comes in, these are integrated in the design of the application.

2.3.5. Testing

The application was tested in a localhost environment. Simultaneous access to the server using mobile devices such as smartphones, tablets, and laptops was conducted.
2.3.6. Implementation
The face detection app was embedded in an emulated ASIST website environment. Next, the randomly selected IT students were asked to access and use the application using smartphones, tablets, and laptops and were asked to evaluate the same later on.

For objective 2, a survey rubric designed by Mutt Susan (2014) to evaluate student applications was adopted. The Student App Review Rubric, features five sections (or criteria) that students can grade when assessing an app. Each of these criteria can be graded with a numerical number from 0 to 4 with 4 as the top grade. According to Jaafar, 2008, usability evaluation provide answers regarding whether the system was designed and developed according to the user’s requirements.

2.4. Data Analysis
In relation to the student and faculty perception of the application, the data gathered were interpreted using the suggested interpretation of the data like a great app is one that gets a score ranging between 20-17, a good one between 16-13, an average 12-10, Needs work 9-6, and a bad app get a score of 5-0.

3. Findings

3.1. Viola-Jones Algorithm
Several algorithms have been used and tested for computer vision on face detection technique, however, the Viola-Jones algorithm for detecting object such as face and has high accuracy rate with 98.38% using real-time video camera[8]. Figure 3 illustrates how Viola-Jones algorithm works in this study.

![Figure 3. Viola-Jones algorithm](image)

The initial step of Viola-Jones algorithm is the Haar-like features that has weak classifiers. Haar-like features are a simple method or tool to describe a certain object within an image. Each Haar-like feature describes the object by describing the difference in its intensities at a certain region. Therefore it can confirm whether an image contains this object by checking if certain regions have differences in their intensities similar to the object to be detected [9]. An example that would clarify this concept is the Haar-like feature describing the region of eyes and cheeks within a human face. This is based on the fact that the eye and eyebrow region have darker intensities than the cheek region. Therefore if the image at this region happens to have a large difference in its intensities, this implies that the image may contain a face.

After finding all the possible Haar-like features that may exist within a window of a given size, each of these features is treated as a weak classifier and is passed on to the AdaBoost algorithm to start training. AdaBoost will also receive training images as an input and will run each given weak classifier on each input image. According to the results the weight of each classifier will be updated, and the algorithm iterates again.

A function to calculate the sums of each rectangle is called integral image. The integral image is a matrix that contains the sums of all pixels on the upper left region to any pixel. This function was tested by generating an image with the darkest level of gray (value =1), and then calculating the integral image. The output Integral Image increases in brightness as it moves towards the right and moves downwards.
since the integral image sums up all the values on the upper left of the current pixel. The final step is a cascading classifiers to reduce the amount of computation and therefore finalize the result of giving a true face or not [10].

3.2. Embedding mother-tongue-based instructional material in image-recognition configuration in Python.

Now, the researchers will present the functions and considerations in the development of the application. First, the general function that manipulates how the face detection behaves relies on the following code snippet based on the tracking.js framework.

```javascript
myTracker.on('track', function(event) {
    if (event.data.length === 0)
    {
        // No targets were detected in this frame.
    } else {
        event.data.forEach(function(data) {
            // Plots the detected targets here.
        });
    }
});
```

In the proposed face detection application, once that the camera detects a face, it activates a line `x.play()` function and do the `x.pause()` function when it cannot. Second, the coding of the face detection was ensured that all viewing devices can be used to access the E-learning material by setting the initial-scale equal to 1. Third, the important libraries that must be included in the face detection application such as tracking.min.js, face.min.js, dat.gui.min.js, state.min.js, bootstrap.mins.js, jquery.min.js, and bootstrap.min.js. These libraries ensures that tracking.js primarily works well.

In the succeeding discussion, we’ll present how student will interact with the application.

![Figure 4. Face Detection trigger button](image)

Figure 4 shows the trigger button embedded in the emulated ASIST website that allows face detection using tracking.js as indicated by a blue down arrow.

![Figure 5. User approval](image)

It can be seen on Figure 5 a dialog box that requires user approval on the use its device camera by the application.
Figure 6. Face detection playing the IM

It can be seen on Figure 6 Face detection playing an IM. The application automatically plays the instructional material once it has detected a frontal-face and pause the same material when it cannot as seen on Figure 7. So, student focus on the material is a requirement in playing the instructional material. This finding corroborates the study of Ugurlu, Y., (2012) that head posture images provides adequate information in increasing human-computer interaction in e-learning system.

Figure 7. Face detection pausing the IM

3.3. The perceptions of the students and faculty members using the student app review rubric.

Figure 8. Student and Faculty App Perception

It can be seen on figure 8 on the student and faculty perceptions. 90% of the students and faculty members perceived the application to be a great app and 20% rated the application to be good. This implies that the students and faculty members agree that the face detection application is great application in controlling an instructional material.
4. Conclusion
Face detection technique can be applied in a multimedia instructional material to add excitement and it can be used to change the method of interaction of the student with an instructional material. Face detection application is great application in controlling an instructional material.

5. Recommendations
The researchers strongly recommend the improvement of this research along the areas of illumination control and face recognition in lieu of face detection. In this case, the learning is delivered to the intended learner before the frontal camera.

References

[1] Barnouti N H, Al-Dabbagn S S M, Matti W E, and Naser M A S 2016. Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance.
[2] Nielsen J B 2010 Face detection and recognition in video-streams Kongens Lyngby 2010 (IMM-B.Sc-2010-14).
[3] Al-Bardi S, and Al-Sawadi N 2014. Face Detection and Recognition for Automating Class Attendance Systems (College of Engineering Electrical and Computer Engineering Department).
[4] Zhang D, Zhou L, Bringgs R, and Nunamaker J F 2005 Instructional Video in E-learning: Assessing the impact of interactive video on learning effectiveness.
[5] Ugurlu Y 2013 Human interactive e-learning systems using head posture images World Academy of Sci., Eng. and Technol. 65 p 201.
[6] Susan M 2014 Digital Divide and Conquer, A Great Student Rubric for Reviewing Apps, Available: http://www.educatorstechnology.com [Accessed: 16 Sept. 2017].
[7] Jaafar A 2008 Malaysian smart school courseware usability study: the effectiveness of analytical evaluation technique compared to empirical study WSEAS Transactions on Information Science and Applications 5(4) pp 342-8.
[8] Chintalapati S and Raghunadh M V 2013 Automated Attendance Management System Based On Face Recognition Algorithms.
[9] Viola P and Jones M J 2004 Robust real-time face detection Int. J. of Computer Vision 57(2) pp 137–54.
[10] Makovetsky R and Petrosyan G 2013 Face detection and tracking with web camera Fundamentals of Computer Vision.