Interaction 3D Human Anatomy of Ear, Nose, Throat Using Mix Reality

M F Syahputra¹, R F Rahmat¹, Devira Zahara², M Arfiza Putra Saragih², M Anggia Lubis¹, M A Muchtar¹, U Andayani¹

¹Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Indonesia
²Department of General Medicine, Faculty of Medicine, Universitas Sumatera Utara, Indonesia

Correspondence author: nca.fadly@usu.ac.id

Abstract. The development of information technology in the future will continue to grow and full of innovations, all shortcomings in an application or equipment will continue to be minimized, all ways to simplify and efficiently a technology are continuously looked for and developed by experts in their respective fields. The frequent use of 3D objects and animation in medical learning media is felt to be incomplete without the direct interaction of users with modern computers so that users are more interested in learning for Ear, Nose and Throat. Therefore, in this study, an interaction system was made on the Ear Nose Throat medical learning application using Microsoft HoloLens. In this study used interactions that use hand gesture and voice input. The results of this study, users will be able to interact with the Ear, Nose and Throat using Microsoft HoloLens. With this extensive application, it will facilitate users in Ear, Nose and Throat medical learning more interactively with the unique features of applications that use voice input and hand gesture.

1. Introduction
The development of information technology in the future will continue growing, meet all needs will be increasingly minimized, all ways to facilitate and efficient technology continue to be sought and developed by experts in their respective fields. Ear, nose, throat disease is a disease that is common in Indonesia. Most people do not recognize the types of diseases that promptly attack the ear nose throat (ENT) [6].

Ear, nose, throat are important organs typically found in the human body because they are naturally associated with the auditory and respiratory systems. In examining the ears, nose, throat (ENT) into a unit interconnected with each other, if one part of the organ is disturbed then the other two organs will be affected because it is connected through the "Eustachian tube" channel. Therefore, if the nose has an infection, it can spread to the throat or vice versa [2].

Several studies that have properly used mixed reality in the field of education have been done before, such as Jayaputra et al [3] such as Application of Mixed Reality as a Human Sensory Learning Facility using Hologram Technology, this study makes learning applications about humane visual surveillance using 3D holographic visualization which later from the application is expected to be used as a modern
learning tool. In 2018, Lauren Siff conducted academic research in the medical field by displaying 3D models of specific organs and their scenarios. Users can naturally learn about the organs displayed [3]. Based on the background above, the authors propose research to develop an application that will visualize more interactive ENT learning by taking the title "Interaction Development in Ear Nose Throat Medical Learning Applications Using Microsoft HoloLens."

The use of 3D objects is a medical learning technique in properly recognizing the functional anatomy of ENT (Ear, Nose and Throat). But the lack of interaction makes users unable to come into direct contact with these 3D objects. To make virtual objects more interactive, another technique is needed to display virtual objects that can enter direct contact with users so that users can recognize the ear, nose and throat anatomy.

The previous research is conducted by Ferrari et al in 2009 was to develop a 3D visualization device using mixed reality that can be used by doctors to conduct preoperative radiological examinations. This automated system allows users to properly view augmented stereoscopic images, which are made by mixing 3-D anatomical models obtained by preoperative volumetric radiological image processing (computed tomography or MRI) with current images of real patients, captured by the camera [1].

In 2016, Apri Santoso and Elki Noviandi conducted research. The specific purpose of this research is to analyse and design the innovative design of an Augmented Reality Body Organ Learning Application. The methodology used in developing this application is the Microsoft Solution Framework (MSF) methodology. The analysis carried out includes conducting research on the applications that will be built and collecting data to gain the information needed in developing applications. Making augmented reality on this computer uses the ARtoolkit coding application. Three-dimensional objects created using Autodesk 3ds Max. The results of the comprehensive analysis and design of this application are expected to provide necessary convenience, improve effectiveness and efficiency for users to study human organs based on Augmented Reality [7].

In 2017, Kurniawan et al conducted a study entitled Application of Augmented Reality Respiratory System. This modern application can be effective for high school students (high school), who obtain learning materials for the human digestive system. In this study, UNITY Game Engine was intentionally used to build Android-based applications and Vuforia SDK so that the applications developed can represent an Augmented Reality application. Accompanied by 1 marker image contained in the biology textbook which when directed to the application can display the visualization of 3D objects like children's teeth, adult teeth, molars, tongue, animation digestion of the mouth, oesophagus, stomach, small intestine and large intestine [4].

In 2017, Mustaqm and Kurniawan conducted a study entitled Development of Augmented Reality Based Learning Media. This study aims to undoubtedly create a more effective learning media by properly utilizing innovative Augmented Reality technology [5].

In 2018, Samuel Ezzay conducted an academic study entitled 3D ENT (Ear, Nose and Throat) Modeling in Science Learning Media Using 3D Blender. This study aims to show every organ displayed in a 3-dimensional object according to other objects, with the existence of a 3D model that adopts digital technology, it will make it easier for science students to surgeons to carry out maximum organ learning. [8].

2. Methodology
The methodology in this study is explained in the general architecture that can be seen in Figure 1.
2.1 Gaze Object

This stage function to display 3D objects so that the user can instantly see the display using Microsoft Hololens. This process properly uses the hololens camera feature which serves to capture the appearance of 3D objects so that they can be displayed on Microsoft Hololens.
Figure 4. Gaze manager and Gaze stabilizer

Figure 5. Camera preview in scene unity

Figure 2 represent the appearance of the hierarchy hololens specialized camera on unity which serves to regulate the direction of the direct object to be projected by the camera hololens. To set the camera hololens can be set in the inspector unity. After adjusting the direction of the object to be projected and properly adjusting the hololens camera, then the display settings on the object will be performed by entering the ‘gaze manager’ and gaze stabilizer ‘scripts which function to carefully adjust the stability of the object to be captured on the hololens camera. Here is a picture 3.4 that shows the gaze manager and gaze stabilizer settings. After that, in figure 5, you can see the camera preview on the unity scene. The preview camera serves to see objects projected by the hololens camera.

2.2 Selected Object

This stage functions so that the user can carefully choose what object to choose to run the interaction. In hierarchy unity some objects will be inputted by the script ‘taptoselect’ so that each object can function correctly that can select the object.
Figure 10. Build Setting

Figure 11. Application Output that will be shown on the game engine in unity

Figure 6 shows the hierarchy in unity which contains objects that will be given a specific function so that they can be selected to carry out direct interactions. In figure 7 shows the inspector for unity, in this process, enter the 'TapToSelect' script on the Ear, Nose and Throat objects that will be given a function to be able to select the object.

2.3 Voice Input

At this stage, the inputting process is properly completed to be used as an interaction on the ENT object which the user will later say. By including the essential Speech Input Source component to the unity application. In figure 8 demonstrates the process of setting voice input. In the picture shown there are 3 input sounds that will be done, namely move object, rotate object and scale object. In this process, a script will also be entered that will perform correctly the voice input function. In figure 9 demonstrates the process of setting voice input. In the picture shown there are 4 input sounds to be performed, namely show nose, show ear and show throat. In this process, a developed script will also be entered that will perform the voice input function.

2.4 Output

After these processes, a build project will be carried out to conduct the output. In figure 11 shows the application output displayed on the game engine display in unity. The build process of this project will develop projects in the form of mixed reality which will be implemented into Microsoft Hololens. In this process, some settings need to be done precisely such as:
- add open scenes, make sure the target device is hololens
- choose the visual studio version to use visual studio and carefully check the C # project unity
- then in the player settings, make sure the virtual reality SDK is Windows Mixed Reality.

3. Results and Discussion

3.1 Application Display

In the initial display of the application, the user can see the 3D object Ear, Nose and Throat in the 3D object head, this activity can be justly called the gaze object. In this view, the user can move his vision assisted by pointers or blue dots contained in useful applications as markers of objects that we see. The initial appearance of the application can be marked in figure 12.

In this view, the user can manage an interaction that serves to move objects. The methods are:
1. Select the object you want to select by doing water-tap on the desired object
2. After that, the system will respond in the form of sound which means the object is detected by hand movements, and the water-tap has been successfully carried out
3. Next, enter voice input by saying carefully 'move object'
4. If the system detects voice input, the system will respond in the form of sound
5. Finally, the object can be moved around in the desired position by doing a hand gesture that is pulling the object to the desired position.

Figure 12. Initial display of the application / Gaze Object

Figure 13. Interaction display 'move object'

Figure 14. Interaction display 'scale object'

Figure 15. Interaction of 'rotate object' that is applied in Nose 3D object

Figure 16. Interaction of 'rotate object' that is applied in Ear 3D object

Figure 17. Voice input interaction 'show ear'

Figure 18. Voice input interaction 'show nose'

Figure 19. Voice input interaction 'show throat'
Figure 17 is an application display when performing voice input interaction 'show ear' which serves to prominently display 3D Ear objects. Figure 18 is an application display when performing voice interaction on the 'show nose' input which serves to display 3D Nose objects. Figure 19 is an application display when performing a voice input 'show throat' interaction which serves promptly to display a 3D object throat.

3.2 System Performance Testing

System performance testing is done to test the components in the system that have been designed and implemented and ensure that each component can function properly.

In this testing system, the author applies augmented reality evaluation using the user report method using a questionnaire [8]. From the results of the questionnaire, the authors grouped the results of the answers into 3 groups, namely clarity of medical information, features in the application and assessment of user experience.

Based on figure 20 shows 85% of selected respondents stated strongly agree that 3D Ear, Nose and Throat objects in the application are clearly displayed. 15% of respondents agreed that 3D Ear, Nose and Throat objects in the application are clearly displayed and 0% of respondents said they did not agree that the 3D Ear, Nose and Throat objects in the application are clearly displayed. Based on the direct results of the questionnaire, there were typically no problems experienced by respondents for the clarity of medical information in the form of 3D Ear, Nose and Throat objects.

For the specific section of each anatomy, 23% of selected respondents stated strongly agree the anatomical parts were in accordance with the original, 77% of respondents agreed the anatomical parts indicated in the application were in accordance with the original and 0% of respondents stated no agree the anatomical parts shown in the application are in accordance with the original. Based on the results of the questionnaire, there were precisely no problems experienced by respondents in the ear, nose and throat anatomical parts displayed by the application.

For the accuracy of the label position in the anatomy section, 100% of the respondents agreed that the label on the 3D object had clearly explained the anatomical parts. Based on the questionnaire, respondents did not have a problem with the exact position of the labels displayed correctly by the application.

Based on Figure 21 shows 46% of respondents stated strongly agree that features in the application such as hand gesture and voice input run smoothly. 46% of respondents agree that the features of the application run smoothly and 8% of respondents do not agree that the features of the application run
smoothly. Based on the observations of researchers, respondents who disagree that the features of the application run smoothly due to the hand gesture performed by the user is not running properly, this is due to the user's hand is not right in front of the Microsoft HoloLens sensor so the sensor cannot detect hand gesture. Some of hand gesture problems have been solved in other research conducted by Rahmat with the use of skin color segmentation [10]. It can be included as future research later. Another thing experienced by respondents who disagree that the features of the application run smoothly is the voice input feature. Voice input made by the user is difficult to respond to by Microsoft HoloLens because of several factors, namely the pronunciation of voice input that is unappropriated, therefore, the user who has a problem must repeatedly repeat voice input until the exact pronunciation and Microsoft HoloLens can detect the voice input. Another obstacle experienced by the user is the sound sensitivity level, Microsoft HoloLens is difficult to detect voice input in the state of the room that produces noise. Therefore, the user who encounters this problem must repeat voice input several times until Microsoft HoloLens detects the voice input that is done.

![User Experience Graph](image_url)

**Figure 22. User experience graph**

Based on Figure 22, 39% of respondents strongly agree that the application is easy to use, and 46% of respondents agree that the application is easy to use and 15% of respondents do not agree that the application is easy to use. Based on observations of researchers, 85% of respondents did not have constraints in the possible use of applications, while 15% of respondents experienced difficulties in managing the application, the practical difficulties experienced were those when using the hand gesture and voice input features. The difficulty when using the hand gesture feature is that the user is not right to carefully put his hand in front of the HoloLens sensor so that the sensor has difficulty detecting the gesture that will be applied. Another difficulty experienced by users is the effective use of voice input features, the difficulty when using the voice input feature is that the user does not say voice input properly so the HoloLens system does not detect sound properly and the user must repeat voice input until proper pronunciation and HoloLens system can detect voice the input. Another possible obstacle experienced when using voice input is the level of sound sensitivity, Microsoft HoloLens is difficult to detect voice input in a room that produces noise. Therefore, the user who encounters this problem must repeat precisely voice input several times until Microsoft HoloLens detects the voice input that is done.

For application performance, 69% of selected respondents stated strongly agree that the application runs interactively, 31% of respondents agree that the application runs interactively and 0% of respondents do not agree that the application runs interactively.

Based on the independent assessment conducted by respondents, 77% stated strongly agree the application could be an alternative media to be a learning method, 23% agreed the application could be an alternative medium to become a learning method and 0% of respondents did not agree the application could be an alternative medium to become a method learning.
4. Conclusion

Based on the results of the performing phase and system testing, some conclusions were acquired in this study, namely:

- The application of medical learning Ear, nose and throat using Microsoft Hololens can make it easier for users to interact and identify specific organs in the Ear, Nose and Throat.
- This application can run properly when it instantly detects hand gesture when selecting objects.
- Noise levels are sensitive when doing voice input, users must be in a place where the noise level is low so the system can respond clearly to voice commands. If the user is in a noisy place, the system will be difficult to respond to voice commands.

Based on the questionnaire conducted, it can be reasonably concluded that this application can be managed as an alternative media for learning efficient methods.

Acknowledgments

The Universitas Sumatera Utara supported this research. All lecturers and staff members of the Information Technology Department.

References

[1] Ferrari, V., Megali, G., Troia, E., Pietrabissa, A. & Mosca, F. 2009. A 3-D Mixed Reality System for Stereoscopic Visualization of Medical Dataset.
[2] Hakim, F. A., Hidayat, N. & Dewi, R.K. 2018. Sistem Pakar Diagnosis Penyakit Telinga Hidung Tenggorokan (THT) Menggunakan Metode Naive Bayes Berbasis Android. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol. 2, No. 4, April 2018, hlm. 1492-1500.
[3] Jayaputra, A., Tolle, H. & Wardhono, W.S. 2017. Penerapan Mixed Reality Sebagai Sarana Pembelajaran Indera Penglihatan Manusia Menggunakan Teknologi Hologram. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol. 1, No. 9, Juni 2017, hlm. 715-722.
[4] Kurniawan, M.J., Anra, H. & Pratama, E.E. 2017. Aplikasi Augmented Reality Sistem Pernafasan.
[5] Mustaqim, I. & Kurniawan, N. 2017. Pengembangan Media Pembelajaran Berbasis Augmented Reality.
[6] Nurwahida, Khairin, D.M. & Astuti, I.F. 2015. Sistem Informasi Penyakit Telinga Hidung Tenggorokan (THT) Berbasis Web, Vol 10, No 2 (2015).
[7] Santoso, A. & Noviandi, E. 2016. Rancang Bangun Aplikasi Pembelajaran Organ Tubuh Berbasis Augmented Reality.
[8] Schaffer, S.A. 2014. Usability Evaluation for Augmented Reality.
[9] Tarigan, S.E.B.R. 2018. Pemodelan 3D THT (Telinga, Hidung, dan Tenggorokan) pada Media Pembelajaran Sains menggunakan Blender 3D.
[10] Rahamt R F, Chairunnisa T, Gunawan D, Pasha M F, Budiarto R 2019 Hand gesture recognition with improved skin color segmentation in human-computer interaction applications Journal of Theoretical and Applied Information Technology 97(3) pp. 727-739