

Interfacing the Microsoft Face API and Kinect for Windows to Create a Personalized, Augmented Reality MRI Mirror

Helena H. Makendengue, Xingnan Xia, and Bennett A. Landman

Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN

KEYWORDS. Augmented reality, Microsoft Kinect, Microsoft Face API, MRI, mirror

BRIEF. An imprinted nanoporous gold substrate was optimized to create a low power Raman biological and chemical sensor.

ABSTRACT. Augmented reality (AR) technology, though highly pervasive, has yet to be developed and implemented in fields of education on a widespread and successful scale, leaving students with less exposure to technology in their classrooms. Previous AR technologies focused on education have been expensive and difficult to develop, leaving instructors with limited teaching resources. In this study the AR magnetic resonance imaging (MRI), an educational, augmented reality program, was created and assessed. The purpose of this program is to allow students to enrich their learning of human anatomy by examining their own anatomy in real time. The AR MRI is capable of predicting a user's age and gender and displaying MRI images of organs that correspond to the user's predicted age and gender in the appropriate locations on their body. After its creation, the AR MRI was assessed on its ability to accurately place the MRI images and detect a user's age and gender. Ultimately, the AR MRI managed to detect a user's age, on average, within ± 6 years of each user's actual age and succeeded in correctly detecting a user's gender 100% of the time. Additionally, the AR MRI demonstrated accurate placement of the various MRI images onto the correct regions of users' bodies with some misalignment.

INTRODUCTION.

As technology becomes increasingly omnipresent in today's world, the need and relevance for learning and interacting with technology is not only beneficial, but also necessary in order to maintain today's technological growth and standards. Because society's interest in augmented reality (AR), in particular, continues to heighten, engineers are striving to find more innovative ways to convey and deliver their ideas to the public. Due to the pervasiveness and flexibility of augmented reality, there are many ways to create inexpensive, yet interactive and highly useful systems, from head mounted displays to mobile phone applications to marker-based applications, which utilize machine-readable labels that are placed within a user's surroundings and recognized by an AR system. Many AR applications that have been produced and implemented worldwide, such as Snapchat and Pokémon Go, provide whimsical, fresh, and innovative ideas to capitalize on recent initiatives in AR, but often fail to provide material for understanding fundamental, educational topics. Google glass has also attempted to pioneer AR technology for education, however it is not without a hefty price tag. The use of a Kinect allows for more inexpensive, personalized alternatives when attempting to create AR programs. The Kinect is a markerless, motion-sensing input device produced by Microsoft with the ability to track six users simultaneously in real-time [1]. It is a relatively common, yet sophisticated system that must be developed in order to serve the desired purpose of being an effective AR learning tool.

This study set out to create and assess an AR device that allowed students to enrich their learning, specifically of the human anatomy, by examining their own anatomy in real time using MRI images. MRI is an imaging technique that utilizes magnetic fields to create images of the internal human anatomy. This study was divided into two phases, 1 and 2, in order to better evaluate and differentiate the design features and creation of the AR system from the functionality of the system. In Phase 1 of this study, a markerless-based, interactive, educational augmented reality program, titled the AR MRI, was created. This program is capable of predicting a user's age and gender and displaying MRI

images that correspond to the user's predicted age and gender in the appropriate locations on their body. Additionally, these MRI images move with the user in real time. The primary goal in Phase 2 of this study was to assess the accuracy of the AR MRI, specifically the accuracy of its age and gender detection features in conjunction with the accuracy at which the system places MRI images onto a user's body.

With AR, students will have the opportunity to further liven and personalize the curriculum that is taught within their classrooms by incorporating digital information and graphics into their surroundings.

MATERIALS AND METHODS.

Phase 1.

During Phase 1 of this study, the AR MRI, an interactive, educational augmented reality program was created. This program allows users to view and interact with digital MRI images of various organs that are overlaid on top of their bodies. The overlaid images that appear superimposed on top of a user move with the user in real time and are personalized based on the user's age and gender. After a user's age and gender is detected, the information is displayed in a text-box above their heads. Once the MRI images appear onscreen, the user can proceed to hover his or her hand over specific images to view cross sections, names, and other details about the organs found within the MRI images.

This system's setup utilizes a PC that runs Microsoft Visual Studio 2015, a Kinect for Windows v2, a 1920 x 1080 projector, and a 96 in. x 56 in. projector screen, all of which are used for the AR Mirror Calibration & Game [2]. Figure 1 illustrates the markerless-based, hardware setup of the AR MRI.

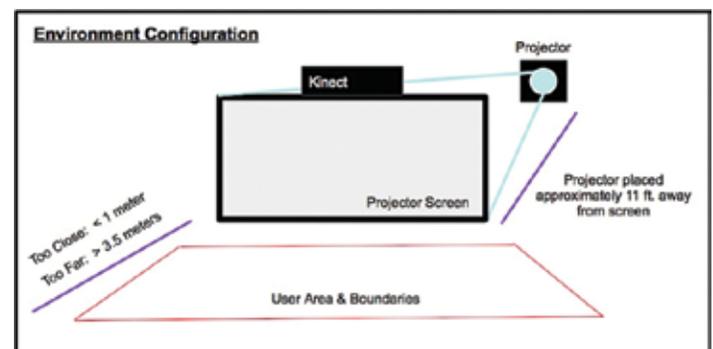


Figure 1. AR MRI Hardware Setup and Environment Configuration.

The construction of the AR MRI began with interfacing the Microsoft Kinect v2. The interfacing of this device allowed the AR MRI to become a markerless-based system, enabling users to incorporate the use of their bodies. The Microsoft Face API is a programming interface with the ability to identify faces in a single photograph and analyze attributes, such as age, gender, facial hair, etc. using various face algorithms. In order for a user to have his or her face analyzed by this API, a photograph of the user must be taken. The AR MRI was designed to take a screenshot of each user's face discreetly when the Kinect detected a body and run the screenshot through the Face API in order to detect and display the user's age and gender [1-3].

Additionally, a JSON database containing all of the MRI images was constructed in order for the AR MRI to pull and display only the images that corresponded with a user's detected age and gender. These images were collected and de-identified from a study at Vanderbilt University led by Dr. Brian Welch. Each image contained its own unique set of properties, such as age, location on the body, gender, etc. The database was interfaced with the age and gender information obtained from the Microsoft Face API, and the matching images were chosen and displayed onto a user's body. The program currently works so that the MRI images are divided and stored into five different age groups. The groups are as follows: 10 years & under, 11-20 years, 21-30 years, 31-50 years and 51+ years.

The tracking, placement, and scaling of the images onto a user's body was done with algorithms that utilized a combination of predefined body points determined by the Kinect and manually defined points. These points are different from markers in the fact that they are programmed to be recognized automatically by the Kinect and do not require machine-readable labels. Ratios were then used to fine-tune and resize the images further. The rotation of images as a user rotates his or her body was also achieved using trigonometric algorithms [1]. Cross section images were also displayed on screen whenever a user had his or her hand hovered over a particular organ shown in the MRI images overlaid on their body. A point for each organ was defined, and when a user's hand was within a radius of X pixels from an organ's point, the image, name, etc. of the cross section for that specific organ appeared on screen.

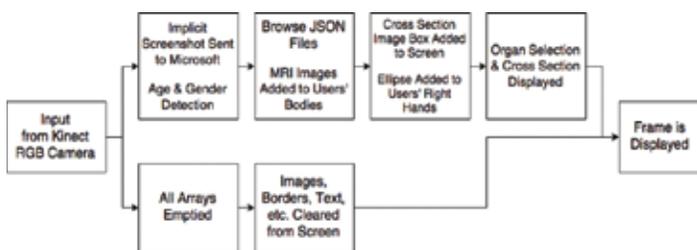


Figure 2. Graphic illustrating how the various devices, APIs, and methods interact with each other to create a functional AR MRI system.

Phase 2.

Concluding the creation of the AR MRI, the functionality of the system was assessed. Participants were gathered to test the AR MRI program, each user ranging in age from 18 to 37 years old. Of the total number of participants, 41% were female, while the remaining 59% participants were male. Each participant was instructed to place five markers on specific parts of their bodies, and these markers were used to determine whether the MRI images that were to be overlaid onto their bodies were positioned accurately. After the placement of the markers, each user was instructed to stand in Kinect's field of view and have their age and gender detected. MRI images of organs that corresponded to their detected age and gender were then overlaid onto their bodies in their designated locations. Screenshots of each user were then taken as he or she interacted with the system. The accuracy of the location of the MRI images was analyzed using these photos. Once every user concluded his or her use of the AR MRI, they exited the view of the Kinect camera and the system was reset.

RESULTS.

Phase 1.

In Phase 1 of this study, a markerless-based, interactive, educational augmented reality program that would allow students to enrich their learning of the human anatomy by examining the uniqueness of their anatomy in real time using MRI images was successfully created. The program utilizes the Kinect for Windows v2 and Microsoft Face API for its complete functionality. It also possesses elements of detecting, tracking, and continually collecting information on its users.

Phase 2.

In Phase 2 of this study, the accuracy of the AR MRI, specifically the accuracy of its age and gender detection features, as well as the accuracy at which the system places MRI images onto a user's body, was assessed.

After the construction of the AR MRI, it was hypothesized that the Kinect's ability to detect certain landmarks on a user's body, in addition to the algorithms that were used in this study, would create accurate, regional placement of the MRI images onto a user's body. During the assessment of the AR MRI, it was found that each organ was placed in the correct region of every user's body, but not every organ was aligned accurately upon placement.

Based on trends noticed while interfacing the Microsoft Face API, it was hypothesized that the API would detect a user's age within 5 years of the user's actual age. However, the average deviation from a user's actual age was calculated to be 5.705 years, or approximately six years. Figure 3 displays each participant's actual age as compared to their detected age. Additionally, it was hypothesized that the AR MRI would correctly detect a user's gender 95% of the time. However, the AR MRI correctly determined the gender of all participants.

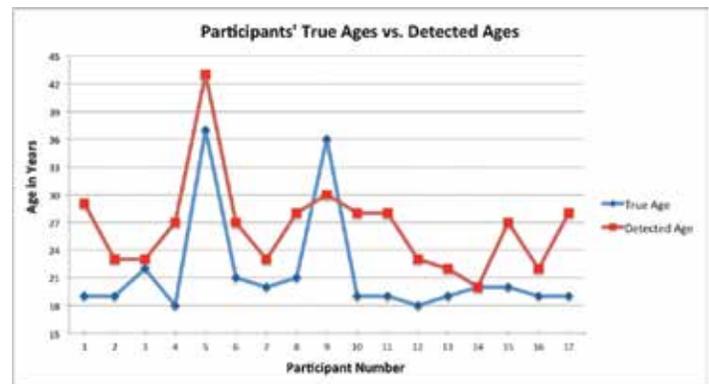


Figure 3. Chart plotting the differences between the participants' actual ages and their ages as detected by the AR MRI.

DISCUSSION.

Phase 1.

A markerless-based, interactive, educational AR program that would allow students to enrich their learning of the human anatomy by examining the uniqueness of their own anatomy in real time using MRI images was successfully created. This program was capable of predicting a user's age and gender and displaying MRI images that corresponded to the user's predicted age and gender in the appropriate locations on the user's body. The incorporation of gender and facial recognition software into the program enables a user's observation of their own anatomy to become more personal, thus contributing to a user's educational experience. The implementation of a database into this program makes possible the observation of various anatomies corresponding to various ages and genders, each with their own unique anatomical structures and characteristics. The ability of the user to view the names and cross sections of the various organs found within the MRI images displayed by hovering their hand over certain areas of their bodies also allows for a more interactive and educational experience. As a result, all design objectives were achieved.

Phase 2.

Despite these objectives having been achieved, each feature implemented spawned, to some degree, a limitation to the accuracy and educational value of the AR MRI. For example, when a user is having his or her face analyzed by the Microsoft Face API, only one photo of the user is collected. Essentially, the API is relying on this photo to accurately represent a user's age and gender. As a result, a user can easily distort age and gender detection results produced by the Microsoft Face API by styling their hair in a more feminine or masculine

manner, shaving facial hair or wearing faux facial hair, or altering their facial expressions. The placement and angle of the Kinect camera, if not setup to view a user from head on at the user's eye level, can also contribute to a distortion in a user's age and gender results. As mentioned previously, the Kinect is only able to track up to six users at a time, with only the first two users identified by the Kinect being tracked in detail. Unfortunately, this limited tracking ability could result in a less accurate alignment of the MRI images onto a user's body. Images in the JSON database must be entered manually, as it is not live or connected to a self-updating database. The program currently works so that the MRI images are divided and stored into five different age groups, however a future goal is to store images for each possible age year. Due to time constraints only a handful of participants were able to test the AR MRI. An IRB for this study also limited eligible participants to 18 years or older. Ideally, with more time, the focus would be K-12 students.



Helena H. Makendengue is a student at Martin Luther King Jr. Academic Magnet High School in Nashville, TN; she participated in the School for Science and Math at Vanderbilt.

CONCLUSION.

More extensive and detailed testing should ensue to more accurately define and adjust the factors that affect the accuracy of the AR MRI system. In addition, further modifications and testing of the algorithms that enable the placement of the MRI images onto the user should take place in order to further improve the accuracy of the image placement. A more encompassing, accurate gender and age detection API should also be used to ensure more accurate gender and age detection, thus improving the accuracy of the images pulled from the JSON database [1]. More textual information and additional organs could also be implemented to expand upon the educational value of this program and its potential as a resource for learning. Additional testing should also take place to assess the potential of the AR MRI program as a learning tool. The AR MRI could be further modified to account for other unique user attributes such as height, weight, and race, creating a hands free experience that blends reality seamlessly with technology.

AR is a versatile form of technology used in a variety of different fields and professions, and AR devices typically vary in complexity and expense. With that noted, this study set out to apply the flexibility and potential of AR to create an AR device that allowed students to enrich their learning of human anatomy by examining their own anatomy in real time. The AR MRI, an interactive, educational AR program was created. It was successful in predicting a user's age and gender and displaying MRI images that corresponded to the user's detected age and gender in their appropriate, general locations on the user's body. With more developed AR technology similar to the AR MRI, society will one day more effectively learn and recall, not just information on the human anatomy, but possibly other disciplines of knowledge as well.

ACKNOWLEDGMENTS.

I would like to thank Dr. Bennett Landman for the opportunity to work in the MASI lab, my mentor, Xingnan Xia, for her guidance and assistance, the School for Science and Math at Vanderbilt and my advisors, Dr. Stephanie Weeden-Wright and Dr. Marci Howdyshell, for their advice and continuous support.

REFERENCES.

1. "WindowsPreview.Kinect Namespace", *Msdn.microsoft.com*, 2016. [Online]. Available: <https://msdn.microsoft.com/en-us/library/windowspreview.kinect.aspx>. [Accessed: 06- Jun- 2016].
2. Yao Yuang, Christopher P. Lee, and Sharo T. Hawrami. "Software Design And Interfaces Implementation For The Augmented Reality Mirror Project."
3. "Microsoft Cognitive Services - Documentation", *Microsoft.com*, 2016. [Online]. Available: <https://www.microsoft.com/cognitive-services/en-us/face-api/documentation/overview>. [Accessed: 28- Jun- 2016].