Deep gesture interaction for augmented anatomy learning

Ahmad Karambakhsh\textsuperscript{a}, Aouaidjia Kamel\textsuperscript{b}, Bin Sheng\textsuperscript{a,\textasteriskcentered}, Ping Li\textsuperscript{b}, Po Yang\textsuperscript{c}, David Dagan Feng\textsuperscript{d}

\textsuperscript{a} Department of Computer Science and Engineering, Shanghai Jiao Tong University, China
\textsuperscript{b} Faculty of Information Technology, Macau University of Science and Technology, Macau
\textsuperscript{c} Department of Computer Science, Liverpool John Moores University, UK
\textsuperscript{d} School of Information Technologies, The University of Sydney, Australia

\section*{ARTICLE INFO}

Keywords:
Neural network
Augmented reality
3D reconstruction
Medical education
Mobile cloud

\section*{ABSTRACT}

Augmented reality is very useful in medical education because of the problem of having body organs in a regular classroom. In this paper, we propose to apply augmented reality to improve the way of teaching in medical schools and institutes. We propose a novel convolutional neural network (CNN) for gesture recognition, which recognizes the human's gestures as a certain instruction. We use augmented reality technology for anatomy learning, which simulates the scenarios where students can learn Anatomy with HoloLens instead of rare specimens. We have used the mesh reconstruction to reconstruct the 3D specimens. A user interface featured augment reality has been designed which fits the common process of anatomy learning. To improve the interaction services, we have applied gestures as an input source and improve the accuracy of gestures recognition by an updated deep convolutional neural network. Our proposed learning method includes many separated train procedures using cloud computing. Each train model and its related inputs have been sent to our cloud and the results are returned to the server. The suggested cloud includes windows and android devices, which are able to install deep convolutional learning libraries. Compared with previous gesture recognition, our approach is not only more accurate but also has more potential for adding new gestures. Furthermore, we have shown that neural networks can be combined with augmented reality as a rising field, and the great potential of augmented reality and neural networks to be employed for medical learning and education systems.

\section{1. Introduction}

In pursuit of immersive human–machine interaction, researchers have explored the different interacting method with new input sources other than the traditional mouse and touchpad. In augmented reality, gesture control is often considered as an ideal interacting method, while leaving gesture recognition as a crucial problem to study (Mitra & Acharya, 2007; Pigou, Dieleman, Kindermans, & Schrauwen, 2015; Qian, Niu, & Yang, 2013; Rautaray & Agrawal, 2015). In contrast to the traditional mouse, the gestures to convey the same instruction are different in every person. Gestures are also a dynamic process so that the duration of the gestures cannot be fixed. Plus, except for the starting and ending positions, every other position in the duration must be detected and track. In the previous work on gesture recognition, gestures cannot be detected and tracked with satisfactory accuracy. Meanwhile, the neural networks have achieved many unprecedented results in deriving meaning and recognizing an object from complicated and vague time in various fields. With this remarkable advantage, neural networks can be used to extract patterns and detect the trends which used to be considered too hard for computers. Thus, a trained neural network can be used to analyze the gestures, such gestures are often more complicated comparing to click the mouse.

Successful and accurate gesture recognition can significantly improve the sense of immersion and user experience. It especially has great potential in medical teaching and learning where require high immersion to simulate the real environment of operations or anatomy courses. Heard from medical school, the professors and students are facing such dilemma that the number of anatomical specimens is limited while the number of students keeps increasing. Students have lacked the opportunity to have a close study of the specimens. Therefore building a three-dimensional object can help understanding which cannot require the textbook.

In this paper, we propose a user interface of augmented anatomy learning with gesture interaction based on the deep convolutional neural network. It offers the functions which teacher can use it to replace the common anatomy teaching process and students can use it to review anytime and anywhere. Also as a tutor, it can be the best choice, for example, medical students by listening to the recorded audio of their
professors and using proposed application can learn anything without joining to those courses directly. Fig. 1 shows the application interfaces of professors and using proposed application can learn anything without joining to those courses directly. Fig. 1 shows the application interfaces displaying human Humersus supporting labels on/off, scale, move, and rotate function by gestures controlling. Students can operate the application by Pan, Pinch, Fist, and Tap gestures which are recognized by a trained deep convolutional neural network. The networks include a 3D convolutional neural network to merge and analyze the information from the depth camera and RGB image. Furthermore, by using a cloud, the procedure of learning is getting faster. In this method, the learning process should divide into some separated server and the results returned back to the main server. Besides the gestures recognition, we have used the mesh reconstruction to 3D reconstruct the specimens. We scanned the model from different view angles and then merge and remesh the scanned meshes by the key points surface representation (KSR) algorithm.

We have shown that neural networks as a rising field can be applied to augmented reality for improvement. We also demonstrate the great potential of augmented reality and neural networks to be employed for medical and educational usage. Before 2012, people mostly use principal component analysis (PCA) to reduce dimension then feed to support vector machine (SVM) to recognize the hand gestures. After 2012, convolutional neural network (CNN) has become an important tool for object recognition since ImageNet of Krizhevsky, Sutskever, and Hinton (2012) have excelled results on the ILSVRC12 challenge. With high-performance GPUs, CNN's show great power on image recognition. Compared with other neural networks, CNN's take fewer parameters with better feature extraction quality which are easier for training. Our architecture contains 26 layers except for the Relu activation. The only sizes of filters here are 1x1 and 3x3. We alternately use these two kinds of filters. All convolutional layers are followed by fully-connected layers. The achievements made by this research include:

- To provide a low-cost and efficient way to reconstruct a 3D model from divided meshes. Scanning objects are done by the rangefinder camera. From the point, scanning of the whole part of a body organ is not possible or at least the quality will decrease. We have divided the scanning into small parts and merged them with KSR method;
- To provide an user-friendly interface, which meets the demands of medical education. The improvement of the user interface is not only gesture recognition, but also a new user interface is designed to interact with an operator in an easy-using manner;

- To provide an efficient way to recognize human's gestures by neural networks. By utilizing the convolutional neural network, the accuracy of the gesture recognition is improved and the operator is able to send a proper command to the augmented reality application.

The remaining parts of the paper are organized as follows: Section 2 is about related work. In Section 3, the details of our approach are presented. In Section 4, the experiments and analysis of our design are presented. The conclusions of the paper go in Section 5.

2. Related work

The origin of human–computer interaction (HCI) to other areas of study such as computer interface design, human factors, usability and specifically to educational environments are examined (Berg, 2000), and now it is a time for progress this way to make it as convenient as possible. For the recent years, many approaches have been proposed for the immersive human–machine interface and augmented reality. We briefly review some studies related to mesh reconstruction, gestures reconstruction, and neural networks.

2.1. Mesh reconstruction

There are lots of methods to reconstruct a mesh model by range sensors or scanners. RealSense camera is one of them that can do the scanning with its SDK and tools. However, the result of the scanners is not acceptable in some fields of usage; also it cannot build object just by one try. Using poison to reconstruct a surface from oriented point samples acquired with 3D range scanners is one of the famous methods in 3D reconstruction, but it runs a risk that the data will be over smooth. Kazhdan, Bolitho, and Hoppe (2006). Calakli and Taubin (2011) made efforts in incorporating positional constraints by using poisson reconstruction algorithm. Furthermore, Kazhdan and Hoppe (2013) proposed screened poison surface reconstruction. It is one of the best surface reconstruction and already implemented in some tools and library such as Meshlab and PCL. Another method to combine all mesh parts together is about using iterative closest point (ICP) algorithm by Holz and Behnke (2015). They had proposed registration of non-uniform density 3D point clouds using approximate surface reconstruction that it can be used to merge all parts from different angles and extract the full object mesh with reasonable quality. The paper that is the main

Fig. 1. General idea to utilize interactive learning based on augmented reality glasses.
دریافت فوری
متن کامل مقاله
امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات