Analysis of the Combination of AR Technology and Translation System

Yuluo Wang[1,*]

[1] *University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA*
[*]*Corresponding author. Email: ywang883@ucsc.edu*

ABSTRACT

Human society has entered the 21st century for two decades already. The rapid development of science and technology is obvious to all. A variety of novel products, including VR technology, are also constantly maturing, coupled with the continuous facilitation of transportation, and the increase in the number of international practitioners. Machine translation technology has become very important. Based on the current existing AR application technology, this article combines artificial intelligence to analyze the possibility of fusion of machine translation and virtual reality. At the same time, it analyzes the existing combination types on the market, such as sign language AR translation, menu AR translation. Combined with the AR glasses product of the current new era, a translation learning technology using AR glasses combined with artificial intelligence big data is proposed. Research has found that there is no perfect product that combines augmented reality with artificial intelligence and machine translation technologies. Due to the huge amount of computation required to generate the virtual world, and the limitations of the current development of AR technology, it will take a long time for such products to integrate into life.

Keywords: *Machine Translation, Augmented Reality, Virtual Reality, Artificial intelligence*

1. INTRODUCTION

Virtual reality is an emerging technology and a derivative of the rapid development of modern computer technology. Due to certain technical problems, virtual reality is always used in science fiction and science fiction movies.. Many subsets are derived from the development of virtual reality, among which augmented reality and mixed reality are the most commonly used. Gradually, Augmented reality is used in life, it penetrates all walks of life and cooperates with it. These include the immersive experience in-game, the visual model in the design field, and the AR Hawk-eye technology in sports competitions to give the fairest judgment and so on. The products of the 21st century are not just AR technology, among which machine translation is also booming. In the 21st century with convenient living and transportation, going abroad is already a breeze, so international trade is constantly increasing, and the most important thing in international trade is communication. It also promotes the demand for a translator. In general, machine translation can be divided into 4 periods, the early exploration boom period(1933-1966), the commercial rule-based period(1967-2007), the statistical machine learning period(1993-2016), and the last period neural network translation period starting in 2013[1]. The easiest way to clearly understand the meaning of the other party is to use a translator. The combination of AR and translation optimizes the current translator on the market and brings opportunities for improvement in translation. This article will provide an introduction to the technical background of AR, the way of combining AR with artificial intelligence, and the combination of AR and translation with existing platforms including local language translation and sign language translation. Moreover the bottleneck of current technology will be discussed, hoping to provide better suggestions and ideas for the development and integration of translation in the future.

2. REASONS TO USE MACHINE TRANSLATOR

The portion of GDP of exported goods has been growing rapidly since 1940 and it has been around 25% recent years, which means there is one-fourth of world trades are international trades[2]. There are countless large international companies such as Apple, Microsoft, etc. The internationalization of trade will become an

inevitable trend in the development of an enterprise. Nevertheless, the language barrier become one of the major issues amount international business. The cost of hiring translators is not low for start-up companies, and they may not be able to accurately translate without specific domain knowledge, and it is difficult to have a neutral translation from the third-party.

Communication is the most important thing in life, but because of different language barriers, it is difficult for people from different regions and countries to communicate if they are not familiar with each other's language. At the same time, deaf-mute people who have lost the ability to express themselves have also lost the opportunity to communicate with this society. For commercial trade, since 1940, the proportion of exported goods in GDP has been increasing rapidly and has been around 25% in recent years, which means that a quarter of world trade is international trade [2]. With the development of economic globalization, major international companies such as Apple and Microsoft continue to increase their overseas markets. Therefore, language barriers become one of the main problems in international business. In the tourism industry , the biggest problem for many major tourism countries when receiving foreign tourists is also the language barrier. Because the demand for translators is still high, but the cost of hiring translators is not low, and if there is no specific domain knowledge, an accurate translation may not be possible, and it is also difficult to have a third-party neutral translation. The emergence of machine translation can help these people regain the ability and opportunity to communicate. People from different countries and languages can communicate directly.

3. EXISTING AR APPLICATION IN TRANSLATION :

3.1 AR + Sign Language

There are many people with disabilities in the world. Among them, deaf-mute people are the most common. In addition to writing, sign language is the most commonly used in their usual communication. According to statistics from the World Federation of the Deaf and Mute, there are more than 72 million deaf people in the world, and there are more than 300 sign language languages nowadays. Due to the particularity of the population, they do not have many educational resources, it's difficult for them to learn sign language. Ordinary translators cannot meet their needs because sign language is a combination of 3D movements. AR sign language translation is a necessary existence for deaf-mute people, who can know how to express a language anytime, anywhere.

AR plus Sign language can also be realized on the technical level. The translation of AR technology plus sign language relies heavily on the database. The certain movement that need to be recognized will be stored in the Vuforia database, and the SDK of the database will be downloaded and imported into Unity3D for programmer to develop[3].

3.2 AR + foreign dish menu

In the modern economy where the economy is very developed, more and more people travel abroad and taste the world's cuisine. The names of many dishes may only be expressed in the local language. For example, the English translations of Chinese glutinous rice balls and classic dumplings called "dumplings". The lack of language translation often prevents people ordering food in foreign restaurants from ordering what they really wants.

At the technical level, the target image can be directly stored in the database according to the characters of different countries. In the research of [4], they used Android phones, and first used the modified fast corner detector implemented using Vuforia Library to quickly Detect the text, and then render a 3d model next to the text. It can be seen from the experimental results that this is feasible, but the requirements for the GPU are relatively high. Using the mobile phone in the experiment to render the five 3d model GPU practicality is as high as 95%.

3.3 AR + rarely used language

The society is gradually becoming unified. Due to educational reasons, children gradually reduce the frequency of speaking local languages. For example, children in many countries have English as their first language, and they gradually forget their native language, that's a cultural lack. They couldn't communicate well with their grandparents who was not able to speak English, which also greatly affected the blood relationship. Therefore, many translations database are needed to record various small languages. So, many historical linguists could use AR translation to pass on the local language, such as BandoAR technology, which is a AR translation application for rarely used language called Banjar[5]. The Banjar language is a good example of inheritance and understanding of minor languages.

4. THE COMBINATION OF AR SYSTEM AND TRANSLATION SYSTEM

4.1 AR

The augmented reality system consists of three parts: a camera, a computational unit, and a display. After capturing an image, the system augments virtual objects on top of the result and displays them. In order to

achieve the enhancement goal, the AR community uses various methods, the most important ones are posing computation and object recognition, those systems are highly dependent on databases. They register objects in the database first and then scan to find them in that database[3]. Every part of the augmented reality system can be extended with many details. In camera part, different ways of scanning on the camera will produce completely different effects. At present, mobile phones are the most commonly used scanning tool, which is usually used to scan text. The most commonly used technology to support this operation is the optical recognition system. Optical character recognition is a very necessary technology in the society. It is used to convert handwritten letters or words into digital format. Optical character recognition technology has many stages, including pre-processing, classification, post-acquisition, pre-processing, segmentation processing, post-processing feature extraction and so on[6]. Each step of optical recognition is important and interrelated, and the most perfect recognition effect can be achieved by improving various technologies at each stage.

4.2 Translation algorithm

After using the OCR system to get digital text, the system needs to translate the text to the target language. Most of the translation software will use the most common machine translation algorithm-neural network. The neural network is to imitate the human brain to get the optimized result layer by layer. The algorithm is very complicated. Using the fully connected neural network as an example, all variables in each adjacent layer will be linked. So, if there is a layer neural network with m parts for each layer and its fully connected neural network, it will have $m^{(n-1)}$ total connections. However, the common used neural network technology is not fully connected, but there is still plenty of calculation within the neural network. After using the neural network translation, the result will be projected onto the surface of the translation object and displayed in different ways on the monitor.

4.3 AR combined with AI to help translation:

AR technology can not only be combined with translation, but also with the most popular artificial intelligence technology. The combination of artificial intelligence and virtual reality and Augmented Reality was originally used to help eliminate inefficiencies in many fields, and for information transmission. Its purpose is to enhance and participate in activities such as game film production and education. Medical students can use artificial intelligence and virtual reality technology to practice simple to complex operations with equipment to improve their skills and reduce risks [7]. It can also be used for patient rehabilitation. The AR technology three-dimensional scan is performed on the patient through artificial intelligence data analysis to obtain the before and after changes to confirm the recovery situation, and also able to use in post-care and pre-care.

The combination of artificial intelligence and augmented reality can also be used in scanning images and acquiring the required environment. Use the effect of fast reading and recognition, which includes automatic text detection technology: when recognizing characters in an image, quickly find one of the characters, and spread to the surroundings to identify the necessary characters in the picture combined with optical character recognition. Then, use the text translation API to find the text in the database.

4.4 Other exists technology that helps AR translation: Speech recognition technology

Real-time speech translation has three components: automatic speech recognition (ASR), machine translation, and speech synthesis [8] The first is to use Automatic speech recognition components to process the original speech. There are many existing technologies that process this speech-to-text. For example, the Microsoft Speech API is built into the Windows system but only supports a small number of languages, and there are also other ASRs such as Sphinx, and Java Speech APIs. Program developers can choose different ASRs according to different needs and different computer languages. The next component is machine translation technology, which translates the text in the original language into the target language. Finally, through the Text to speech component, the translated target text language is synthesized into speech reading text.

5. DISCUSSION

The use of machine translation is very extensive. There are more than 500 million people are using Google Translate, which is equivalent to 1.5 times the population of the United States. The mainstream people who use machine translation are students and scholars studying abroad, or people who do business abroad. For those who live in the United States but are not professional with English, using phone based translator is particularly troublesome to take out a mobile phone and computer to translate. Using translation technology to combine with the existing augmented reality glasses is a good solution. Based on the current rapid development of technology, AR glasses have become cheaper and cheaper, slowly entering people's lives. In the second half of 2019, Vuzix, a leading company in the field of AR (augmented reality) equipment, and Zoi Meet, a voice captioning service company, plan to provide real-time translation services on Vuzix Blade

AR smart glasses. The Zoi Meet application can translate 12 languages in real time, which means that anyone who uses Vuzix Blade AR glasses can communicate with others through real-time translation. This is a good product for many people who have communication needs with a foreigner but are not able to speak other languages.

Combining AR and translation can also better help students and scholars studying abroad to learn well. When reading an article in a second language, it often takes a lot of time to input the characters into a computer. Also, inputting will make the study become noncontinuous and disrupt the thinking. In real-time translation glasses, use optical character recognition to capture dynamic text, then use VR plus artificial intelligence technology to identify and calculate the subject category that is currently being studied, and use big data to compare and compare the types commonly used by users to make an optimized translation. Finally, appear on the user's AR glasses with optimal translation. That allow users to save a lot of time and energy during studying period.

6. TECHNICAL LIMITATIONS

6.1 Technical problems and drawbacks of Machine Translation

The idea of using machine translation language was put forward 50 years ago. With the advancement of technology and databases, translation technology has gradually improved and more people have used it. But machine translation is still challenging because natural language is highly vague. In other words, two languages often express the same content in different ways, including grammatical phrases[8]. The current level of machine translation is still not strong enough to replace the level of oral translation. For the kind of advanced articles and content with high accuracy requirements, machine translation is too prone to errors and problems. There are also many immature aspects of translation technology. For example, people from different regions and countries in the google map will make reviews on places of interest. For many emotional words, the translation algorithm uses less active English words[9].

6.2 The shortcomings of AR glasses

In the current development market of AR glasses, there are many exciting models and styles, such as uzix M100, Google Glass,and Microsoft Hololens, etc. However, these products are very few in daily use. Even though such products are technically helpful for people's life, in the public view, most people are unwilling to wear AR glasses mainly due to size, weight and appearance. Nevertheless, from the research of Elisa Maria Klose and Ludger Schmidt, it can be found that

people who are willing to use AR glasses are more inclined to the effect and function like the display technology, ocularity than appearance[10].

7. CONCLUSION

The whole article discusses the importance of translation technology in today's society, outlines the combined presentation of AR and different types of translation, analyzes the translation technology AR technology and artificial intelligence technology, and proposes a combination of AR smart glasses and translation technology. Combine the concept and do some analysis on it. However, many technologies are still developing, there is no perfect product for the technologies in the fields of augmented reality combined with artificial intelligence and machine translation. However, with the development of technology for more than half a century and the recent rise of neural network, it has been able to be similar to human level and even surpass human. In the current popularity of big data, if translation technology can be combined with big data analysis to make its own different translation effects and results for different industries and fields, it must be an invention for the benefit of mankind.

The first commercial application of AR appeared in 2008, and the development and popularization of computers has just passed half a century. In the future, with the improvement of computer science and technology and the emergence of new commercial applications, AR technology for commercial purposes will continue to develop and update. Due to the lack of specific programming information and related data content, this article is mainly inclined to analyze the current AR technology-related applications. With the continuous development of AR software and hardware, some experiments such as making the AR glass thicker or make user more immersed while using it can be completed in the future.

REFERENCES

[1] Algorithm Engineer, Artificial Intelligence. (n.d.). A brief history of machine translation. Zhihu special column. Retrieved September 17, 2021, from https://zhuanlan.zhihu.com/p/223033781.

[2] Beltekian, E. O.-O. and D. (n.d.). Trade and globalization. Our World in Data. Retrieved September 17, 2021, from https://ourworldindata.org/trade-and-globalization.

[3] N. Soogund and M. H. Joseph, "SignAR: A Sign Language Translator Application with Augmented Reality using Text and Image Recognition," 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal

Processing (INCOS), 2019, pp. 1-5, doi: 10.1109/INCOS45849.2019.8951322.

[4] D. Arioputra and C. H. Lin, "Mobile augmented reality as a Chinese menu translator," 2015 IEEE International Conference on Consumer Electronics - Taiwan, 2015, pp. 7-8, doi: 10.1109/ICCE-TW.2015.7217035.

[5] A. A. Syahidi, H. Tolle, A. A. Supianto and K. Arai, "BandoAR: Real-Time Text Based Detection System Using Augmented Reality for Media Translator Banjar Language to Indonesian with Smartphone," 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS), 2018, pp. 1-6, doi: 10.1109/ICETAS.2018.8629251.

[6] A. M. Sabu and A. S. Das, "A Survey on various Optical Character Recognition Techniques," 2018 Conference on Emerging Devices and Smart Systems (ICEDSS), 2018, pp. 152-155, doi: 10.1109/ICEDSS.2018.8544323.

[7] J. Kiruthika and S. Khaddaj, "Impact and Challenges of Using of Virtual Reality & Artificial Intelligence in Businesses," 2017 16th International Symposium on Distributed Computing and Applications to Business, Engineering and Science (DCABES), 2017, pp. 165-168, doi: 10.1109/DCABES.2017.43.

[8] T. Duarte, R. Prikladnicki, F. Calefato and F. Lanubile, "Speech Recognition for Voice-Based Machine Translation," in IEEE Software, vol. 31, no. 1, pp. 26-31, Jan.-Feb. 2014, doi: 10.1109/MS.2014.14.

[9] B. Mathayomchan and K. Sripanidkulchai, "Utilizing Google Translated Reviews from Google Maps in Sentiment Analysis for Phuket Tourist Attractions," 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2019, pp. 260-265, doi: 10.1109/JCSSE.2019.8864150.

[10] E. M. Klose and L. Schmidt, "A User-Based Comparison of Two Augmented Reality Glasses," 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 2018, pp. 1-2, doi: 10.1109/VR.2018.8446175.

[11] Javornik, A. (2016, October 4). The mainstreaming of augmented reality: A brief history. Harvard Business Review. Retrieved September 29, 2021, from https://hbr.org/2016/10/the-mainstreaming-of-augmented-reality-a-brief-history.