AN INTELLIGENT MOBILE AR SYSTEM BASED ON NEURAL NETWORK

SUNG-BONG JANG

Department of Industry-Academy Cooperation, Kumoh National Institute of Technology, South Korea
E-mail: sungbong.jang@kumoh.ac.kr

Abstract - An intelligent mobile AR system dynamically combines the most appropriate object with the input video to minimize the user's intervention. In the system, one of problems is how to choose the most appropriate object from huge number of AR objects. To solve the problem, this paper proposes a method based on a neural network.

Index Terms - Mobile augmented reality, real-time object augmentation, fully connected neural network

I. INTRODUCTION

The fourth revolution is attracting much interests to the people after it has been mentioned in the 2016 world-wide economic forum by Klaus Schwab. The core technology of the IR 4.0 is the artificial intelligence (AI) that will be used in all of industrial fields in the future. AIs are rapidly evolved thanks to the machine learning. The machine learning is that the system or machine acquire the knowledge by learning and make decision about something by using the knowledge. Originally, learning was the peculiar ability of the human being. However, this ability is being imported to the machine. The machine learning can be categorized into supervised learning and unsupervised learning. In supervised learning, the training data and answers are given to the machine at the same time. As the name implies, the machine are supervised by human being for learning. On the other hand, in unsupervised learning, the answers are not given. Therefore, this kind of machined learning is used to classify some data, not prediction. Recently, the machine learning has evolved rapidly into deep learning. In deep learning, they use convolutional neural network, multi-layer perceptron neural network, and so on. Especially, the rapid improvement of the CPU processing power due to the parallel processing has enhanced the performance of the neural network algorithm. For example, alpha-go developed by Google is being popular for various industrial areas.

Furthermore, Augmented Reality (AR) is one of the important driving technologies that promote the fourth industrial revolution. Especially, the mobile AR can be efficiently used to get various information from the other sites or places. There are three types of mobile augmented reality technologies: vision-based, tag-based, GPS-based. Sometimes, each technology are mixed to implement a useful service. There has been active researches about mobile augmented reality. In recent years, many researches have been done in the literature of mobile augmented reality. I. B. C. Ramon et al. [01] presented a framework that enable creation of marker-based mobile augmented reality applications in mobile devices. The framework consists of five major blocks. The first is a presentation block which visualize the user interface of the mobile AR applications. The second is a tracking block that provides all necessary vision processing functions. The third is a interaction block that provides a interaction function with the users through a graphical editor. The fourth block provides the users with the real representation of the reality easily. The final is a context block. The block provides the system with the environmental situation around the mobile device. C. Dimitris et al. [02] describe about the state-of-the-art of the mobile augmented reality researches. Their descriptions have been done from the perspective of the target application, UI/UX, system compositions, AR object characteristics, network connection, and so on. Especially, they point that in the near future, the cloud computing shall play an important role in booming up the mobile augmented reality. Y. Xin et al. [03] presented a new algorithm that detecting an object from input video data. The proposed algorithm shows faster execution time and higher performance than the existing image descriptors. S. M. Lehman et al. [04] presented a solution to solve the privacy problems of the mobile AR systems. The application programmers can handle the privacy-threating situation using context data like locations, network status, or other inputs.

II. PROPSED SYSTEMS

This paper proposes an intelligent mobile AR systems. The most important point from traditional AR system is that the most optimal object can be chosen automatically and dynamically combined with the input video. In the existing AR system, the object to be augmented is fixed and not changed. However, in the proposed approach, the object is different from situation to situation. The overall system architecture

is illustrated in Figure 1. The whole system consists of the four number of software blocks: video quality enhancement, real-time object detection, intelligent object augmentation, and frame renderer blocks. Here, this paper focus on intelligent object augmentation block. The other blocks already has been discussed in the previous works. The basic idea of the proposed approach is that the object is dynamically combined with the input video according to the context information that is the environmental situation at the site where the customer visits. The procedure to take the optimal object is as follows. First, intelligent object augmentation block (IOAB) analyze the current video, user's location, and situation, and sends the analysis results information to the AR Objects Analysis Server. The contents of the sent message may be as follows. It can contain user's location, detected objects from input video, around building name, and video id. Second, the AR Objects Analysis Server starts choosing the optimal object based on neural networks. Basically, this works use a fully connected neural networks.

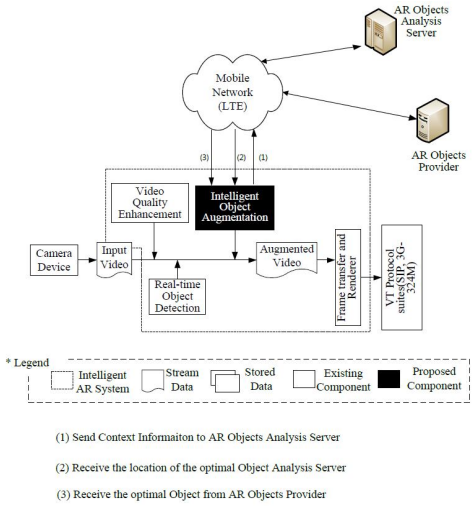


Figure 1. An Intelligent AR System

The machine learning technology is the main driving force of the artificial intelligence development. It can be categorized into two kinds: supervised learning and unsupervised learning. In supervised learning, the answer data are given to the model. On the other hand, for unsupervised learning, there is no target data. There are a number of algorithms that can be used for AI.

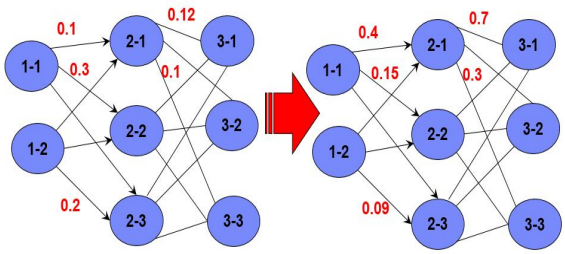


Figure 2 Changes of the weighted values in Fully Connected Networks

The most widely used one is a neural network. The network has three layers: input, hidden, and output layers. Input layer has many nodes which pass the input data, and output layer has the nodes to process the output data. The values to go to next layer are calculated based on the weighted averaged summation. The weights of each edge changes whenever it is learned using training data. Figure 2 represents the weight values of each edge before and after training. The next figure represents a process of defining network and training.

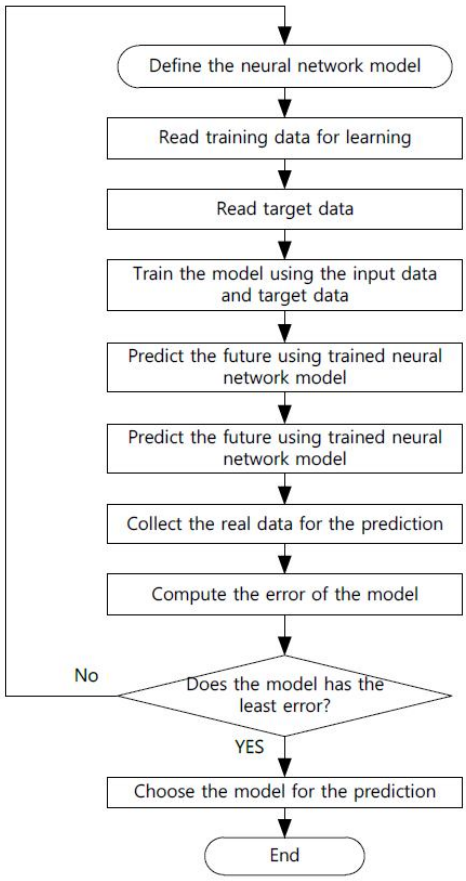


Figure 3 Training Process of the Neural Network for Choosing an Optimal AR Object

The process for choosing the best neural network model is illustrated in Figure 2. First, we define the default neural network model. In this definition, we define the number of nodes for each layer, and define the basic parameters such as default weighted value of each edge. For example, we define the number of nodes as 5, 5*2, 10 for input, hidden, and output. The neural network can have multiple number of hidden layers. However, for simplicity, we define only one hidden layer. In step (2), NN system read the training data for machine learning. Here, input training is defined as follows.

(1) Detected object kind: This represents what kind of object being captured through the user's mobile phone. This may be determined by the analysis of object

analysis and detection.
(2) Detected object's metadata : After analyzing, we can obtain the metadata of the object.
(3) User's location: This represents where the user is located. For example, he may visit book store or shoes store

In step (3), the system read the target data for machine leraning. Here, the target data can be obtained through the historical data of the AR object. To do this, for some time, the system may have to collect the data through a Metadata Server. In step (4), using input and target data, the neural network system must be trained. It will be better if the repetition number is as many as possible. After training, the weighted values of each edge will be changed. This will lead to enhance the correctness of the prediction. In step (5), we have to predict the future using trained neural network. Here, the prediction is done to check the error of the model. In step (6), we collect the real answers for some period. In step (7), we compare the predicted values and real answers to computer the error of the defined neural network model. The most widely used metric for evaluating the model is average absolute error. It is defined as follows.

$$nn_error = \frac{1}{m}\sum_{i=1}^{m}\frac{|\mathrm{Real_val}_i - \mathrm{Predict_val}_i|}{\mathrm{Real_values}_i}\times 100(\%) \quad (1)$$

Here, real_values$_i$ represents a value which has been observed. Predic_val$_i$ represents a value that has been predicted for each values. m is the number of input nodes. Using this equation, error table for each model is constructed. The table I shows an example of error table.

Model	Input Node Number	Hidden Node Number	Output Node Number	Error
Model 1	5	5*2	5	8.3915
Mode 2	15	15*3	10	1.4421
Model 3	5	10	10	1.8980
Model 4	5	15	5	2.3191
Model 5	20	25	10	2.3192
Model 6	30	50	20	1.8369

Table 1. An Example of Neural Network Model Evaluation

As seen from this table, Model 2 has the least value of the error. Therefore, we choose a model 2 for choosing the best AR object. In step (8), we have to check the defined model has the least error or not. If the model has the least error, we choose the model as the final neural network model to be used to choose the best AR Object in an intelligent AR system.

CONCLUSIONS

An intelligent real-time mobile AR can provide the users with more efficient information because the system find the most appropriate object using neural networks. For example, suppose that a user visit a store to buy an electronic device, USB Disk. In the store, huge number of devices are displayed and he was embarrassed, and he needs some additional information for each product to compare the products. In the old department store, he would have to call for a clerk to hear the basic specifications of each product. Therefore, it takes time and cost to get required information for choosing a product. However, if he use a mobile AR adopting the proposed approaches, he can save much time and cost. In the proposed system, the user capture the product using mobile phone camera. Then, the system analyze the input video, send the analysis results to the server, and instantly display the price, and detailed specifications of the disk on the mobile phone. He can get the necessary information as soon as possible automatically. Traditionally, if he want get the information of the product, he must capture the QR code of the product. Now, he does not need to do the action.

ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2016R1C1B1014346). This paper was supported by Research Fund, Kumoh National Institute of Technology.

REFERENCES

[1] I. B. C. Ramon, O. V. V. Osslan, G. C. S. Vianey, "A Mobile Augmented Reality Framework Based on Reusable Components", IEEE Latin America Trans., vol.13, no.3, pp.713-720, 2015.
[2] C. Dimitris, B. Carlos, H. Zhanpeng, H.Pan, "Mobile Augmented Reality Survey: From Where We Are to Where We Go", IEEE Access vol.5, pp. 6917-6950, 2017.
[3] Y. Xin, K.-T. Cheng, "Learning Optimized Local Difference Binaries for Scalable Augmented Reality on Mobile Devices", IEEE Trans. on Vis. and Comp. Graphics, vol.20, no.6, pp.852-865, 2014.
[4] S. M. Lehman, C. C. Tan., "PrivacyManager: An access control framework for mobile augmented reality applications", 2017 IEEE Conf. on Comm. and Netw. Sec. (CNS), pp.1-9, 2017.

★★★