Exploration of the Application of Virtual Reality and Internet of Things in Film and Television Production Mode

Qian Song 1,* and Yoo Sang Wook 2

1 Department of Marine Convergence Design, Pukyong National University, Busan 608737, Korea
2 Industrial Design Department, Pukyong National University, Busan 608737, Korea; yooswook@pknu.ac.kr
* Correspondence: songqian719@gmail.com

Received: 30 April 2020; Accepted: 14 May 2020; Published: 16 May 2020

Abstract: In order to reduce some of the problems of technological restructuring and insufficient expansion in the current film and television production mode, the application of emerging technologies such as artificial intelligence (AI), virtual reality (VR), and Internet of Things (IoT) in the film and television industry is introduced in this research. First, a topical crawler tool was constructed to grab relevant texts about “AI”, “VR”, and “IoT” crossover “film and television”, and the grasping accuracy rate and recall rate of this tool were compared. Then, based on the extracted text, the data of recent development in related fields were extracted. The AdaBoost algorithm was used to improve the BP (Back Propagation) neural network (BPNN). This model was used to predict the future development scale of related fields. Finally, a virtual character interaction system based on IoT-sensor technology was built and its performance was tested. The results showed that the topical crawler tool constructed in this study had higher recall rate and accuracy than other tools, and a total of 188 texts related to AI, VR, and IoT crossover television films were selected after Naive Bayes classification. In addition, the error of the BPNN prediction model based on the AdaBoost algorithm was less than 20%, and it can effectively predict the future development scale of AI and other fields. In addition, the virtual character interaction system based on IoT technology constructed in this study has a high motion recognition rate, produces a strong sense of immersion among users, and can realize real-time capture and imitation of character movements. In a word, the field of AI and VR crossover film and television has great development prospects in the future. Therefore, the application of IoT technology in building the virtual-character interaction system can improve the effect of VR or AI film and television production.

Keywords: AI; VR; IoT; film and television production; virtual character interaction system

1. Introduction

There has been rapid development of computer, virtual reality (VR), and Internet of Things (IoT) technologies against the background of artificial intelligence. VR technology has gone through a series of updates and iterations, and now consumer VR products suitable for the market are appearing in front of people. Now VR technology has been applied in all walks of life, which drives the supporting upgrading of a series of VR industry chains such as hardware, technology, content, and platform; this trend also affects the VR industry in China [1,2]. In 2015, Lighter Chaser Animation and Letin VR Digital Technology both carried out VR film production practice successively. Among them, Letin VR Digital Technology opened the chapter of VR filmmaking in China by making the VR film Live to the Last. However, due to the limited profit channels at that time, the producers could only recover the production cost of the film by shooting advertisements on the video platform and cooperating with the
VR experience pavilion. However, with the development of VR technology and the improvement in people's living standards, VR technology has made a qualitative leap in the field of film and television production since its development [3]. According to data, by 2021, China will have 85 million VR headsets in use, and the revenue from VR content will reach over 3.5 billion US dollars [4]. Meanwhile, some cinemas in China are ramping up investment. Erdong Pictures plans to set up a virtual reality cinema in Beijing, which will be the third VR cinema in China. In addition to releasing its own VR microfilms, Digital Domain, a visual effects company, has also set up its own VR studios in China [5]. Although there is still a long way to go in the development of VR film and television in China, the future development of VR film and television in China can be predicted.

When VR technology is used for film and television production, the appearance of virtual characters in films and television dramas is not entirely obtained by pure computer technology. In order to make the movements and expressions of virtual characters more real, sensor technology is often needed to capture the movements and expressions of real characters [6]. Foreign films, such as Avatar, Rise of the Planet of the Apes, Beauty and the Beast, and the Chinese domestic film Legend of Ravaging Dynasties capture actors’ movements and expressions by employing sensor technology, and then make virtual characters similar to actors’ movements and expressions and put them into virtual scenes. Such production methods can help the actors complete difficult movements, make the virtual characters more real, and increase the viewer’s sense of substitution [7]. Sensor technology is the main driver in the IoT; according to the estimation of relevant institutions, driven by the Internet of Things, the compound growth rate of the global sensor market from 2016 to 2021 is estimated to be 11%, and the market size will reach $200 billion in 2021 [8].

In terms of the current production methods that combine VR and IoT with movies and TV, there are still many problems to be solved compared with ordinary movies. In order to explore the development status and future development direction of VR film and television under the background of artificial intelligence, and the application of sensor technology in VR film and television production, based on the development status of film and television production of VR and IoT technology in China, a prediction algorithm for future development trend is proposed based on the development data. Meanwhile, a method for virtual character interaction is proposed based on IoT sensor technology. Finally, the problems in the development and application of this field are pointed out and corresponding solutions are put forward.

2. Literature Review

AI, IoT, and VR are emerging information technologies. They have been gradually applied in various fields, and some scholars have conducted relevant research. The integration of AI, IoT, and VR technologies with film and television production can change the limitations of film and television production, greatly improve the ability of film and television production, and present a shocking visual experience for the audience.

2.1. The Application and Development of VR Technology in Film and Television

There are many researches on VR technology at home and abroad, but most of them focus on the industrial application and commercial practice of VR technology. The research of VR technology has also attracted extensive attention from scholars. The application of VR technology in video production, combined with medical treatment, can improve the treatment of patients. However, there are still many problems in the promotion and application of VR technology. With the continuous development of VR technology and the improvement of people’s living standards, VR technology has also been applied to film and television production. Jones and Dawkins (2018) took the original 360-degree film of Chungking Mansions in Hong Kong as the basis of their research, and combined it with interactive, immersive, and narrative VR. Through comparison of the changes in heat and smell in the viewing experience, it is found that it could stimulate the audience’s senses and lead to increased sense of presence [9]. According to Zhang et al. (2019), the realistic real-time simulation of fluid animation has

Applied Sci. 2020, 10, 3450 2 of 16
been widely applied in practice (such as VR and AR), while fast simulation requires a lot of physical calculation and time-step size [10]. Yu (2017), however, indicated that making ultra-high-quality content was the next major challenge for the VR industry; creating virtual reality, where the human eye was indistinguishable from the real world, would require light-field technology—3D imaging. When computer vision and machine learning were combined, the production cost of optical field technology could be reduced, which has become a feasible way to improve VR quality [11]. Xu and Ragan (2019) evaluated the application of VR technology in film production and proposed the application of virtual characters in travel guide introductions, finding that the application of VR technology to the production of travel-related videos could effectively attract the attention of the audience [12]. In order to improve the sense of immersion in VR films, Wang (2017) combed out the technological process of VR film sound production from the perspective of the audience, including early recording, post-editing and mixing, sound export, and master production [13]. In conclusion, VR technology is widely used in the film and television field. VR-based film and television present a stronger sense of substitution, and the film and television drama produced using VR is more real.

2.2. The Application and Development of IoT Technology in Film and Television

IoT refers to the acquisition of any object or process that needs to be monitored, connected, or interacted with in real time through different information sensors and other technologies. All kinds of information of the target objects are collected, and the connection between objects, and between objects and people is realized through network access, and the intelligent perception, recognition, and management of objects and processes are achieved [14]. Today, IoT technology has been widely used in agriculture, medical, military, and other fields, and IoT technology is also being used in the field of film and television production. Hebing et al. (2019) used sensor technology for motion steps and combined acoustic processing and visualization technology for MovieScape creation [15]. Yu et al. (2019) proposed a fast method to capture human motion based on an RGB-D (Red Green Blue - Depth) sensor, indicating that the method could improve the accuracy and robustness of the capture [16]. Takahashi et al. (2019) proposed a 3D model-construction method of human motion based on a VR system and a multi-camera unlabeled motion capture system [17]. Protopapadakis et al. (2018) performed motion capture of dancers based on Kinect sensor and constructed a human bone data recognition method. It was found that this method could be used to identify multiple postures and obtain accurate body joint information [18]. It can be found that the current research mainly focuses on the security of video data during IoT transmission.

To sum up, researchers have conducted in-depth research on the application of AI, IoT, and VR technologies in various fields, while few researchers have combined these three technologies with film and television. Based on previous studies, in this study, IoT, VR, and AI technologies were integrated, and the application and influence in the film and television industry is discussed.

3. Proposed Method

3.1. Video Data Crawling Based on Theme Crawler

The traditional theme-crawling strategy can be used to analyze the whole content of web pages and determine the relevance of candidate links. In this study, Dewey decimal classification was used to classify keywords in web pages. The keywords used in this study were VR, and film and television, and the corresponding Dewey classification numbers were 505, and 303, respectively. Then, the candidate link topic edge text was used to extract keywords with similar word meanings. The main steps of the extraction process were as follows: Firstly, segmentation processing was carried out on the text of the web page and anchor text, and the vocabulary not used anymore was removed; secondly, WebDewey2.0 was used to search the keyword Dewey classification number; thirdly, the feature extraction of the edge text of the feature candidate link topic was carried out using two-dimensional coordinates combined with Dewey decimal classification, and the two-dimensional coordinates were
drawn; fourthly, according to the distribution trend of key words in 2D coordinates, key words of anchor text key points and their surrounding gathering points were extracted. Then, the Naive Bayes Swinburne classifier was used to analyze the keywords of the theme edge text. Assuming the total number of keywords in the page body and anchor text in the edge text of candidate link theme is \( f \), then Equation (1) can be obtained.

\[
f = f_{\text{body}} + 5f_{\text{anchor}}
\]

In Equation (1), \( f_{\text{body}} \) is the number of keywords in the body of the page and \( f_{\text{anchor}} \) is the number of keywords in the anchor text.

The word frequency of the feature word \( t \) is as follows:

\[
f(t) = f_{\text{body}}(t) + 5f_{\text{anchor}}(t)
\]

Then the weight of the feature word \( t \) is as follows:

\[
V(t) = \frac{f(t)}{f}
\]

Then the Bernoulli model algorithm is used to classify the feature attribute vectors, and the probability of whether the content \( d_i \) in the text belongs to category \( c_j \) is also calculated.

\[
C_{\text{NB}} = P\left( c_j \right) \propto P\left( \frac{d}{c_j} \right)
\]

\[
\propto P\left( c_j \right) P\left( t_{i,1}, t_{i,2}, \ldots, t_{i,n} / c_j \right)
\]

\[
\propto P\left( c_j \right) \prod_{n=1}^{m} P\left( t_{i,n} / c_j \right)
\]

In Equation (4), \( C_{\text{NB}} \) represents the probability weight in naive Bayes Swinburne classifier, and the highest category is the category in the candidate link topic edge text.

The probability values \( P(c_j) \) and \( P(t_{i,n} / c_j) \) in Equation (4) are estimated as follows:

\[
\begin{cases}
P(c_j) = \frac{N(C=c_j)}{N} \\
P(t_{i,n} / c_j) = \frac{N(t_{i,n}, C=c_j) + 1}{N(C=c_j) + M}
\end{cases}
\]

In Equation (5), \( C \) is the text category in the corpus, \( N \) is the number of texts in the corpus, \( N(C=c_j) \) is the number of texts belonging to \( c_j \) in the corpus, \( N(t_{i,n} / C=c_j) \) is the text quantity of the feature words \( t_{i,n} \) belonging to \( c_j \), and \( M \) is the number of keywords in the text at the edge of the topic to be accessed.

The theme web crawler was then used to crawl the keywords in the webpage, and the main workflow was as follows. I. link fetching—the seed links are grabbed in the queue in the initial state (in this study, artificial method was used to obtain seed links); II. The topic-focused crawler requests the network to download the corresponding web page according to the link, after the server receives the HTTP request protocol, the response is returned; III. The obtained webpage is added to the reply queue to wait for the page preprocessing and relevance determination; IV. The topic edge text in the candidate link is identified and extracted; V. The naive Bayer Swinburne classifier is used to determine the topic category of candidate linked topic text, and the text classification corpus is the complete version of Sogou text classification corpus, and the corpus contains eight related topics, such as film and television, artificial intelligence, and virtual reality; VI. The corresponding webpage is put into the database of the background page related to the topic; VII. candidate links are sorted according to the weights calculated by the naive Bayes classifier. In the integration scheme proposed and designed in this study, the self-built film and television database based on Internet of Things and MySQL (My Structured Query Language) were used as the server of the database, and the information extraction rules are shown in the Table 1.
The hardware configuration of this study was 4 G memory, 250 G hard disk, 10 M network bandwidth, and Windows 7 (64-bit) operating system. In this study, the test of the theme web crawler method was proposed using the search accuracy (topic-focused crawlers downloading theme-related web pages/the total number of webpage pages) and the recall rate (topic-focused crawlers downloading theme-related web pages/the number of theme-related web pages in the network).

3.2. Prediction of the Development Trend of VR Film and Television based on the AdaBoost-BP Algorithm

In order to better predict the development trend of VR movies, the AdaBoost algorithm was combined with BP neural network to form the AdaBoost-BP algorithm. The training process of BPNN is shown in Figure 1 and was mainly the initialization of network and parameters. Then, the training set data were input in the input layer, the output results of hidden layer and output layer were calculated, and the error value between the output result and the expected result was calculated. When the model error reached the expected result and the network calculation iteration times reached the maximum set value, the calculation could be completed; if the requirement was not met, the error in hidden layer was recalculated, the error gradient was obtained and the parameter values in the network were updated, and the input calculation for the training set was carried out again.

![Figure 1. The calculation flow-chart of the BP neural network (BPNN).](image-url)

In this study, the AdaBoost algorithm was used in the improvement of the BP neural network (BPNN). The improved prediction model structure is shown in Figure 2. The main calculation flow of the algorithm is as follows. Firstly, the training sample was selected, the BPNN was initialized, and
the weights and thresholds in the BPNN were adjusted according to the training samples. Secondly, the training samples were used to train the weak learning machine, and the prediction error and the distribution weight of the next weak learning machine were calculated. Thirdly, the distribution weight of training samples was adjusted and the weight value of weak learning machine was calculated. Fourthly, the strong learning machine prediction results were output.

![Diagram of the AdaBoost-BP Algorithm](image)

**Figure 2.** The calculating process of the AdaBoost-BP Algorithm.

According to the theme web crawler, the data with VR movie labels were obtained. A total of 8004 data features were included in the 188 relevant texts. In this study, 1000 data were randomly selected as the test set and the remaining 7004 data were selected as the training set. In Matlab software environment, LibSVM toolbox and Matlab neural network toolbox were used to simulate the AdaBoost-BP algorithm.

### 3.3. The Interactive Application-Implementation-Based IoT Technology

In this study, during modeling, sensors were set at each joint and major muscle group of the subjects to collect the movements of the characters. Then, Poser software was used to construct 3D human model, and the obtained human model was exported to the data file in obj format. After writing the program, the information of points, planes, normal vectors, and textures in the data file were read. Finally, Open Graphics Library was used to draw the human model. In the OBJ file, the prefix v represented the vertex coordinates, vt represented the texture coordinates, vn represented the normal vector coordinates, and f represented the integer type surface. When obj data were read, the main steps were as follows. First, read the raw data; second, judged the information of points, planes, vectors and textures; third, recorded the number of information; and fourth, stored these variables and returned to step I.

In the data collection of human bone points, the Kinect device was mainly used. The device contained bone coordinate points of 25 human limbs. The device mainly took the front direction as the Z axis, the vertical direction as the Y axis, the left and right directions as the X axis, and the waist
midpoint as the origin for the construction of human coordinate system. Then the coordinate system obtained by the device was converted to the user’s own coordinate system.

Then, in this study, the distance between SpineMid and SpineBase bone was kept at 0.3, so as to obtain the scaling ratio S. The continuous movement could be regarded as the continuous static posture of multiple sequences, while the static posture mainly included the relationship between bone points, relative positions, relative vector information, and angle information. Therefore, it was necessary to make full use of all kinds of description information in the construction of static posture to accurately describe the object’s posture, and the relationship between different postures is shown in Figure 3.

Based on Figure 3, the relative position relation of the three bone points of the right hand was defined. The right hand coordinate W, x > the right hand elbow coordinate E, x; the right elbow coordinate E, y < the right shoulder coordinate S, y; the right hand coordinate W, z > the right shoulder coordinate S, z. Ten people were invited to conduct the test of motion recognition, including eight boys and two girls, with an average height of 1.74 ± 5.26 m. The subjects all waved, rolled, and squatted, raised their hands, and stepped, and each movement was repeated three times. Subsequently, the accuracy of the proposed method in different motion recognition was compared in this study, and then the difference of recognition accuracy time between the proposed method in this study and other methods was compared.

Based on the development system of Unity3D engine, a set of interactive command actions in 3D space were defined using a Kinect motion-sensing interaction device, which was used to make a naked eye 3D interactive experience system with multiple viewpoints. The specific system framework is shown in Figure 4.

**Figure 3.** Right hand waving in different positions.

**Figure 4.** The basic framework of a naked eye 3D interactive system.
The virtual scene in this system was mainly constructed by 3Dmax modeling. The motion capture of subjects was realized by Kinect sensor and SDK development kit. Moreover, bone binding was used to map the captured motion information to the virtual character, so as to realize real-time control of the virtual character. In order to test the effect of the system constructed in this study, five myopic and five non-myopic people aged from 22 to 26 years old were selected to score the stereo effect, clarity, comfort, naturalness, and immersion of the system, with a total score of 25 (5 for each).

Then, based on the Unity engine, the VR virtual test scene was built. The specific structure of the system is shown in Figure 5.

Immersive interaction and display were mainly personal 3D human–computer interaction, which realized the natural performance of visual display channel through head-mounted HMD (Holographic Media Disk) interactive equipment. In order to better highlight the immersive interactive experience of the system designed in this study, the differences in interaction time and experience scores between mouse operation, Oculus headset, and Cardboard VR headset were selected for comparison.

4. Results

4.1. Validation of Crawler Tools and Prediction Models

Firstly, with the theme of “film and television” as the main topic, the Wikipedia film and television and Baidu encyclopedia film and television were extracted as the seed links in a manual way, and the quality of crawlers in this study was evaluated. A total of 5000 web pages were downloaded by this method, and the recall and accuracy of downloaded text were calculated for every 5000 downloads. The results were compared with best first search and Naive Bayes, respectively, as shown in Figure 6. As concluded from Figure 6A, the recall rate of the study construction method was significantly higher than that of the Best first search and Naive Bayes method; as the number of pages crawled increased, the recall rate of the Best first search method was gradually higher than that of the Naive Bayes method. As concluded from Figure 6B, the accuracy of Best first search and Naive Bayes is lower than the method used in this study, and with the increase of page grasping, the accuracy of the algorithm decreases, but Best first search is the lowest.
The results are shown in Tables 2 and 3. The error values of root mean squared for different weak learning machines were all lower than 20%, while the accuracy error of the AdaBoost-BP algorithm was all lower than 0.5, while the accuracy error of the AdaBoost-BP algorithm as the structure of the subsequent AdaBoost-BP algorithm for the experiment.

Furthermore, the classification accuracy of the algorithm constructed in this study also gradually became stable. Therefore, the number of six weak learning machines was selected as the structure of the subsequent AdaBoost-BP algorithm for the experiment.

The existing data were used to analyze the prediction accuracy of the AdaBoost-BP algorithm. The results are shown in Tables 2 and 3. The error values of root mean squared for different weak learning machines were all lower than 0.5, while the accuracy error of the AdaBoost-BP algorithm constructed in this study was all lower than 20%.
Table 2. Prediction results and weights of each weak learning machine in the AdaBoost-BP algorithm.

<table>
<thead>
<tr>
<th>Weak Learning Machine ID</th>
<th>Root Mean Squared Error</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.35</td>
<td>0.11</td>
</tr>
<tr>
<td>2</td>
<td>0.39</td>
<td>0.10</td>
</tr>
<tr>
<td>3</td>
<td>0.38</td>
<td>0.10</td>
</tr>
<tr>
<td>4</td>
<td>0.42</td>
<td>0.10</td>
</tr>
<tr>
<td>5</td>
<td>0.39</td>
<td>0.99</td>
</tr>
<tr>
<td>6</td>
<td>0.41</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Table 3. Prediction results and error values of the AdaBoost-BP algorithm.

<table>
<thead>
<tr>
<th></th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
<th>Sample 4</th>
<th>Sample 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected values</td>
<td>0.11</td>
<td>0.17</td>
<td>0.09</td>
<td>0.13</td>
<td>0.16</td>
</tr>
<tr>
<td>Predicted values</td>
<td>0.09</td>
<td>0.15</td>
<td>0.07</td>
<td>0.12</td>
<td>0.11</td>
</tr>
<tr>
<td>Error values</td>
<td>0.15</td>
<td>0.14</td>
<td>0.09</td>
<td>0.06</td>
<td>0.18</td>
</tr>
</tbody>
</table>

4.2. Analysis of External Feature Set Prediction Results of VR Film and Television Data

According to the theme crawler, 188 relevant texts about VR film and television and IoT film and television were obtained. The 188 texts were mainly from the state administration of press, publication, radio, film, and television, www.baogao.com, www.qianzhan.com, etc. Among them, 121 were about “VR+ film and television”, accounting for 64.36% of the total text and 67 articles on “IoT+ film and television”, accounting for 35.64% of the total text. Based on the text obtained by the crawling tool, the AdaBoost-BP algorithm was used to predict the development of VR and IoT in China.

Firstly, the development status of Chinese film and television production was analyzed. As shown in Figure 8A, China’s total box office revenue from 2014 to 2019 shows a trend of gradual increase. The total box office revenue of Chinese films increased to 64,266 billion yuan by 2019. As shown in Figure 8B, 9,708 new screens were added in China in 2019, bringing the total number of screens to 69,787 and consolidating the global-leading position.

Figure 8. (A,B) The development trend of the Chinese film and television industry.

Under the background of artificial intelligence, VR technology and sensor technology were increasingly applied in film and television production. As shown in Figure 9, with the arrival of 5G, the current VR industry entered a high-speed development stage (2018). The product form basically took shape, and the user portrait gradually became clear. At this stage, with the emergence of hot style products, the VR industry is expected to further stimulate the market demand for VR products. At the same time, VR applications are expected to be more abundant and interpenetrate with mobile communication, media, education, and other industries, so as to promote the rapid development of the industry.
As an important part of video, film is an important segment of VR content development. In particular, in 2015, the box office of Chinese films exceeded 40 billion yuan for the first time, and the film industry achieved leapfrog development in China. The virtual-reality movie experience is regarded as the advanced version of 3D movie experience, and the development space and speed of VR movie market can be predicted by referring to the speed of 3D development. From the popularity of 3D films in China in 2003 to present, 3D films have maintained an important share. Although the share declined in the initial stage, with the maturity of 3D technology and the participation and interaction of the domestic market, especially in recent years, 3D films have maintained a high proportion in domestic cinema ticket rooms. In 2015, they contributed nearly 50%. Figure 11A shows the growth of VR video equipment, and more and more users are increasingly interested in VR video products. It can be concluded from Figure 11B that the VR live broadcasting market also shows a trend of increasing year by year, indicating that the VR film industry is expected to emerge in the future.
As shown in Figure 12, the market size of Chinese sensors was about 99.5 billion yuan in 2015, and it reached about 160 billion yuan in 2019. The predicted results showed that the sensor scale in China will exceed 200 billion yuan in 2021. Sensor technology based on motion and expression capture is also gradually applied in VR film production. With the gradual expansion of China’s sensor market and the continuous development of VR technology, China’s AI film and television industry will show a trend of rapid development in the future.

4.3. The Application Verification of VR Human–Computer Interaction Based on IoT Technology

In order to explore the practical application of IoT technology in VR film and television production, sensors were used to capture the movement of characters and construct virtual characters to realize VR human–computer interaction. Firstly, the differences in accuracy of the system constructed in this study in different motion recognition was compared. As shown in Table 4 the recognition accuracy of the system varied greatly in different movements, with the recognition rate of tilt reaching 100%, while the recognition rate of wave was the lowest (90.26%).

Table 4. The recognition accuracy comparison of different movements.

<table>
<thead>
<tr>
<th>Motions</th>
<th>Wave</th>
<th>Leaning</th>
<th>Squat</th>
<th>Raising Hands</th>
<th>Stepping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition accuracy</td>
<td>90.26%</td>
<td>100.00%</td>
<td>98.17%</td>
<td>98.26%</td>
<td>90.87%</td>
</tr>
</tbody>
</table>

The system was then compared with other recognition algorithms. As shown in Table 5, the DTW (Dynamic Time Warping) algorithm has the highest recognition accuracy (95.97%), while the video stream matching recognition accuracy is the lowest (86.37%). The proposed method has the fastest recognition speed, does not require offline training in advance, and shows good scalability.
Table 5. Performance comparison of different recognition algorithms; DTW (Dynamic Time Warping), FSM (Frequent Subgraph Mining)

<table>
<thead>
<tr>
<th>Projects</th>
<th>Video Stream Template Matching</th>
<th>DTW</th>
<th>FSM</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition accuracy</td>
<td>86.37%</td>
<td>95.97%</td>
<td>95.33%</td>
<td>91.37%</td>
</tr>
<tr>
<td>Speed</td>
<td>Slow</td>
<td>Relatively fast</td>
<td>Relatively fast</td>
<td>Fast</td>
</tr>
<tr>
<td>Off-line training</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Expansibility</td>
<td>Bad</td>
<td>Relatively bad</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Adapter</td>
<td>2D, simple</td>
<td>Complex, continuous, standard</td>
<td>Simple, continuous, definition</td>
<td>Simple, continuous/non-continuous</td>
</tr>
</tbody>
</table>

Subsequently, the performance differences of different interactive systems were compared. The results are shown in Figure 13A. The average mouse operation time was the shortest (45.80 s), followed by Oculus (46.83 s), and Cardboard VR (60.63 s). As shown in Figure 13B, the scores of sense of reality (1), sense of immersion (2), sense of naturalness (3), convenience (4), interest (5), and satisfaction (7) are all the highest for Oculus, but the acceptability (6) of Oculus is the lowest. After comparing the average scores, it is found that Oculus has the highest average score (4.61), Cardboard VR (4.34), and mouse operation has the lowest average score (4.19).

![Figure 13. (A,B) Performance testing of VR interactive systems.](image)

Finally, it can be found from Figure 14A,B that the virtual character modeling system based on IoT sensor technology constructed in this study can capture the real time movement of the character and map it to the display screen. As shown in Figure 14B, the system can complete the real-time addition of virtual scene. If the subject is wearing a head display device, they can watch the virtual character image constructed by their own motion capture in real time.

![Figure 14. (A,B) IoT sensor-based virtual character construction.](image)
5. Discussion

Some film and television sectors can be equipped with AI or VR technology for better effects. VR technology is relatively easy to carry, and the low-cost column could quickly become the new favorite of VR industry after the reform. VR and AI technologies can bring new ideas to people and make people more interested in film, which can help them seize the market opportunities [19]. The pursuit of visual effects and immersion is a key development project in the era of AI and VR. For some very exquisite pictures and shots, even without the use of AI or VR technology, the audience can feel the contents of TV or movies. However, people are full of expectations for the VR of these column types, and they are sure that this is the most exciting column category in the era of VR. In this study, relevant data were crawled by constructing theme web crawlers. Compared with other methods, it was found that the accuracy and recall of web crawlers constructed in this study were significantly higher than the Best first search [20] and Naive Bayes [21] methods. In addition, there was application of sensor technology in IoT in virtual character construction and VR human–computer interaction. The results showed that sensor technology based on the action steps can realize the establishment of a VR interactive system, which provides a theoretical reference for the production of VR film and television. Based on the relevant texts and data obtained by crawling, it was found that the development prospect of VR film and television in China in the future is very broad. However, there are still many obstacles in VR film production. From the perspective of the shooting and production of creators, the production of AI or VR films is different from traditional films, especially in production, as when people get too close to the camera, they feel “crossed”. Although the later splicing software can splice and combine the videos captured by multiple cameras to a certain extent, it still cannot completely solve the parallax problem [22]. It has been found that the application of motion capture sensors in the motion capture of characters can effectively complete the construction of virtual characters. Moreover, the IoT sensor market in China is also showing a trend of gradual expansion. Therefore, the application of sensor technology in VR film and television production in the future will greatly increase the sense of reality of virtual characters and give viewers a better sense of immersion and interaction, which is conducive to the development of VR film and television.

At present, AI and VR technologies are still relatively immature, which can lead to increased vertigo when viewing [23]. With the development of AI and VR technologies, movie-goers still need to wear head-mounted display devices to watch movies. The strong sense of substitution makes the audience feel that they are present in the film. If the viewing time is still 90 minutes, like traditional film and television works, the audience will not have a good experience when watching movies and will feel dizzy and uncomfortable [24]. Therefore, the length of AI or VR films is usually limited to about 10 minutes. It has been found that, compared with the mouse operation and Cardboard VR system, the construction of a VR interactive system using Oculus has a better sense of immersion and comfort, which provides a reference for solving the discomfort caused by wearing VR equipment for a long time. From the perspective of a head display device, the audience can watch AI or VR movies and TV programs in the head display device, which is an immersive interactive experience with strong privacy [25]. This is different from the experience of watching movies in a group in traditional cinemas and results in higher requirements for the movie mode of AI or VR movies. In a traditional cinema, there is more than one screening room. Usually, a screening hall can accommodate about 100 people, because the cost of a movie can enable more than 100 people to watch a movie together. The movie mode settings for AI or VR movies are different. If 100 moviegoers come to the cinema, the same number of headphones and accessories, such as various sensory simulators, are required [26]. The results of this study indicate that VR film and television will become the mainstream market in China’s VR market segment, so in the future there will be special VR cinemas in China, and this situation will be greatly improved.
6. Conclusions

Compared with traditional film and television, VR film and television are a new panoramic experience, which can attract the audience to experience VR works in a more targeted way. Through the combination of VR technology with film and television, the data can be detected and continuously traced via the sensor of the IoT, enabling users to try immersive experiences in vision, hearing, and touch [27]. The development status of VR film production and IoT sensor technology in China was analyzed based on the AI film production text obtained by the theme crawler tool. The results showed that China’s VR film and television market segment occupies the largest proportion, and sensor technology also has a broad development prospect [28]. A VR film production method under the construction of virtual characters based on motion capture sensor is proposed. It was found that the Cardboard VR system can be applied to the construction of VR interactive system for better integration and comfort. Based on the existing problems in Chinese VR film and television production, the proposed method will be of great significance in solving such problems as difficulties in VR film and television shooting and poor induction of head-display equipment [29]. Due to the lack of specific text rounds, only relevant texts were crawled through network software, while the industry elites in the field of artificial intelligence were not contacted directly for substantive discussion. This was a regret in the process of making the research. However, completing this article was a meaningful challenge, and the creation and research of AI movies and TV programs in China are still in the development stage, so a lot of people are looking forward to seeing what will happen in the future.

Author Contributions: Conceptualization, Q.S. and Y.S.W.; writing—original draft preparation, Q.S.; writing—review and editing, Y.S.W.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We acknowledge the support given from the Pukyong National University.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).