Research Article
Virtual Reality and Its Application for Producing TV Programs

Liya Lin

Fuzhou University of International Studies and Trade, Fuzhou, Fujian 350202, China

Correspondence should be addressed to Liya Lin; linliya@fzfu.edu.cn

Received 24 April 2022; Revised 29 May 2022; Accepted 11 June 2022; Published 28 June 2022

Academic Editor: Abid Yahya

Copyright © 2022 Liya Lin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper aims to conduct an in-depth study on the effective application of virtual reality technology in TV program production. To begin, construct the three-dimensional animation system’s overall structure, determine the animation system’s user interface, and establish the three-dimensional animation system for TV program production. Second, software like YOLOv3 and ResNet-50 network is used for identification and facial recognition. Then, the pixel cross-support window is built using the adaptive stereo matching method. The adaptive filter window is built using the horizontal expansion of the cross-support window. For the successful application of virtual reality technology in TV program creation, the aggregation is completed by regional filtering, and the final parallax image is obtained by parallax selection and parallax optimization. This technology has the ability to improve the visual effect and quality of TV programs, increase the amount of information in programs’ content, and shorten the production cycle of the program in the context of the gradual development of science and technology and the new technology period. In comparison to other approaches, simulation experiments demonstrate that the proposed method has the potential to increase the quality and attractiveness of TV programs and meet the psychological demand of viewers.

1. Introduction

With the increase in the development of science and technology, network data traffic has increased rapidly. Internet technology has brought a scientific and technological revolution. All the laws of the traditional industrial era have been reformed. The rapid development of traditional TV media in the future world is a critical problem that the TV industry practitioners must consider and address. New technology also brings new opportunities and development to the traditional TV media based on great impact. The effective application of virtual reality technology in the TV industry has the potential to open a new world for the TV industry. It is not only technological innovation and change but also a significant transformation and innovation in program planning, design, and production concepts. The traditional TV media is about to turn a new page.

Compared to classic movies, the current production techniques combine VR (virtual reality) and IoT (Internet of things) with movies and TV. However, there are still numerous challenges that need to be solved. Based on the current level of VR and IoT technology in film and television production, a prediction algorithm for future development trends is proposed. Meanwhile, a virtual character interaction system is being developed using IoT sensor technology. Finally, problems encountered in the development and implementation of this sector are addressed, and remedies are given.

The content VR for TV Program creation can be of two types, 360-degree videos & 3D animations. Virtual reality technology can imitate the natural educational environment. Students can utilize the network to access the virtual classroom at any time and from any location. Virtual studio technology is growing rapidly in the TV industry. The new technology of using virtual reality in television can deliver new visual experiences to the audience as well as new tools and create new thinking for television programs. In reality, the use of virtual reality technology in TV program production takes artistic design as the core and technical means as the main carrier to express artistic ideas and create new TV programs. The process of technological realization is incorporated with creative thoughts, and the two belong to an inseparable unity.

The innovations of this paper are as follows:
(1) Face detection in TV programs based on YOLOv3 and face recognition is shown using the ResNet-50 network. The pixel cross-support window is built using the adaptive matching method. The final parallax image is created using parallax selection and parallax optimization.

(2) Simulation experiments reveal that the proposed strategy can successfully improve the quality and appeal of TV programs and meet the psychological needs of the audience.

The rest of the paper is organized as follows. Section 2 explains the related work of different researchers who have already proposed different models based on virtual reality. Section 3 sheds light on the application of 3D animation system structure. Section 4 describes the application of virtual reality technology in TV program production in which face recognition based on YOLOv3 and 4.2 stereo virtual reality technology in TV program production to shorten the production cycle of TV shows, reduce costs, and provide a high-quality viewing experience for the public and provide numerous forms of interaction. TV programs can adapt to the viewing habits of the public in the new media era and better promote the rapid development of the TV industry. This paper discusses the relevant optimization strategies of TV program production and editing but does not provide practical measures to improve the quality of TV programs. Xu proposes a stereo image synthesis method based on multiview images for TV program production. The function curve characteristics of multiview image and camera response are considered, and a virtual multiview image rendering method is then proposed. Different exposure images are drawn from the same viewpoint. To preserve a lot of details and structure, it is necessary to fill the holes and repair the edges of the virtual exposure image. The edge difference is then used to process the edge information of the image. The virtual image is synthesized and used in TV program production. The experimental results reveal that the similarity between the exposure image and the reference viewpoint image drawn by this method is high, although this method has the problem of poor section appeal.

2. Related Work

The rapid development of computer and virtual reality technology makes the implementation of this technology in TV program production logical. The research results have also been achieved simultaneously. According to Wu, the television industry has also grown rapidly with time. The demand for TV programming has gradually expanded with the increase of entertainment in public social life. There are many types of television programming content. To improve the overall competitiveness of TV programs, we must have strong postediting as an effective support to package the program content in order to improve the broad appeal of TV programs and attract the public. The postproduction of TV programs is a comprehensive and technical integration work. It is necessary to carry out comprehensive scientific creation of sound, text, and images and increase the amount of information in the content. The above processing operation is the final phase of TV program production. The level of editing technology will have a direct impact on broadcasting. Therefore, the future technology of TV programs is extensively examined, but no specific methods for improving the overall quality of TV programs are suggested [1]. Wan investigated the digital era and noticed that significant changes have occurred in the life of the public. It has been enhanced by the rapid popularization and development of electronic devices. Multimedia is an important support for the production of digital TV programs. The use of multimedia technology in the production of TV programs is a traditional media innovation. In the production process of digital TV programs, we must comprehensively consider the factors such as TV program quality, production cost, and timeliness. Multimedia technology can be used in program production to shorten the production cycle of TV shows, reduce costs, and provide a high-quality viewing experience to the public. However, this method does not effectively improve the quality of TV programs [2]. Zhao proposed that with the rapid development of new media, the production form of TV programs is gradually evolving to the live broadcast mode. It pays more attention to vision and hearing and primarily emphasizes filtering news materials and paying attention to the relevant work of production, editing, and recreation. As a result, optimizing TV program production and editing can give a better viewing experience for the public and provide various forms of interaction. TV programs can adapt to the viewing habits of the public in the new media era and better promote the rapid development of the TV industry. This paper discusses the relevant optimization strategies of TV program production and editing but does not provide practical measures to improve the quality of TV programs [3]. Xu proposes a stereo image synthesis method based on multiview images for TV program production. The function curve characteristics of multiview image and camera response are considered, and a virtual multiview image rendering method is then proposed. Different exposure images are drawn from the same viewpoint. To preserve a lot of details and structure, it is necessary to fill the holes and repair the edges of the virtual exposure image. The edge difference is then used to process the edge information of the image. The virtual image is synthesized and used in TV program production. The experimental results reveal that the similarity between the exposure image and the reference viewpoint image drawn by this method is high, although this method has the problem of poor section appeal [4].

3. Application of 3D Animation System Structure

The three-dimensional animation system is used in TV program production. The network topology of the three-dimensional animation system is shown in Figure 1.

In Figure 1, the LAN is mainly used by 3D animation designers and relevant producers. The site must be equipped with high-performance computers and graphics workstations. These sites are the main role of the rendering server in network rendering. The renderer is regarded as the manager of rendering. When rendering the network, it is responsible for allocating rendering tasks to other rendering servers and operates the task of rendering works [5, 6].

The structure diagram of the management terminal system of three-dimensional animation is shown in Figure 2. Based on the 3D animation production process, the 3D animation system can be separated into the management subsystem and the production terminal system. The system structure and function of the management terminal include the creative creation scheme in the preparation process of a TV program and the management and distribution of design-related documents. The production terminal aims to solve the challenges in the postproduction of 3D animation [7, 8].

The user interface of 3D animation system can be divided into a three-dimensional animation management interface and a production interface. Figure 3 shows the management interface. The management end of Figure 3 primarily feeds creative, copy, and design requirements documents into the
database of three-dimensional animation programs. Then, the TV program file is sent to the management server. After that, each user who makes short films receives them before sending, managing, and cleaning up the relevant file saved in the file server of the 3D animation at the production end [9, 10]. Analysis and monitoring of the three-dimensional animation programs completed by the server are done at the production end. The qualified ones are added to the television program database of three-dimensional animation, and a request is sent to the examination system of the supporting platform of television program production. If qualified, they are output to other subsystem file storage or broadcast control in the database containing television shows.

The production terminal system primarily consists of various production technologies of three-dimensional animation to complete the creation of different three-dimensional animation materials. It creates various materials for 3D animation and requires the division of operators’ labor. The technology of the production side has a wide range of specialized modeling methods. Modeling can also be divided into the scene, role, and architectural modeling. Different types of modeling technologies can be classified into different sites. These sites can be assigned to character animation, rendering technology, and animation technology. Operators need to be familiar with one site of three-dimensional animation technology. All users use the LAN to assist and cooperate with communication and technology and develop animation systems to create animation materials [11, 12]. In the production of television shows, animation and rendering technology is a widespread and vital technology. In order to produce animation programs, it is necessary to use a motion capture system and rendering factory as an animation production auxiliary system, especially rendering technology. The rendering manager can render materials quickly, while large materials must be rendered by LAN composed of the rendering manager and rendering server to realize the creation of 3D animation materials.
4. Application of Virtual Reality Technology in TV Program Production

Each technology used in film and television creates a different visual experience. VR technology is being introduced into film and television. VR virtual technology in film and television art can partially meet people’s need for a high-quality existence, and it is also a major driving factor in the development of film and television. To improve the high-quality existence of virtual reality, face recognition is used.

4.1. Face Recognition Based on YOLOv3. It is an object detection algorithm that detects specific items in films, live feeds, or photos in real time. YOLO uses features learned by a deep convolutional neural network to detect an item. The face in the TV program is recognized initially when virtual reality technology is used in TV program production. The template box is set in the way of dimensional clustering to predict the bounding box. The network is utilized to predict the relevant variables \((tx, ty, tw, th)\).

\[
\begin{align*}
    b_x &= \sigma(t_x) + c_x, \\
    b_y &= \sigma(t_y) + c_y, \\
    b_w &= p_w e^{b_w}, \\
    b_h &= p_h e^{b_h}.
\end{align*}
\]

(1)

In the training process of bounding box prediction, square and error losses are used to carry out supervision training. The corresponding components of the bounding box are assessed using logistic regression after the result of the bounding box. The model box with the highest score is selected for the effective prediction based on the score to reduce the calculation time.

The virtual model completes the corresponding training in the wider face data set. During the training, the target category, the confidence, and the central coordinates of the target position are supervised by the mean binary cross-entropy, which can be expressed as

\[
\text{loss} = \frac{-1}{n} \sum_x [y_1 n(a) + (1 - y)1 n(1 - a)].
\]

(2)

Photographic robot face recognition is a problem of open-set face recognition, and the main essence is metric learning [13, 14]. After the face detection of the TV program is completed, it must detect and identify the target face, whether the target is the primary character or a secondary character in the program, because, in real work, all the faces cannot appear in the training set. This paper takes ResNet-50 as the basic framework to improve it and performs collaborative supervision in the program production and training process using Softmax loss and center loss to make virtual model recognition separable and compact in class.

The loss of Softmax is expressed by the following formula (3):

\[
L_s = - \sum_{i=1}^{m} \log \frac{e^{W_j x_i + b_j}}{\sum_{j=1}^{n} e^{W_j x_i + b_j}},
\]

(3)

where \(x_i \in \mathbb{R}^d\) represents the characteristics of the \(i\) depth and belongs to the \(y_i\) category. \(W_j \in \mathbb{R}^d\) represents the \(j\) column of the connection layer weight \(W \in \mathbb{R}^{d \times n}\) of the last
layer, and $b \in \mathbb{R}^n$ represents the offset term. $m$ represents the smallest batch, and $n$ represents the number of categories. The learned program production features can be split under the supervision of the Softmax loss function, but there will be significant intraclass differences. By defining the loss function of the center, changes inside the class are minimized in order to improve the separability between classes by assuring the separability of different features. The loss function of the center is defined as

$$L_C = \frac{1}{2} \sum_{i=1}^{m} \left\| x_i - c_{yi} \right\|_2^2,$$

where $c_{yi}$ represents the category center of the $y_i$ program depth feature, which is modified with the change of the program depth feature. However, if we take the update feature center of the overall training set into account in the iteration center, it will increase the amount of calculation and implementation difficulty. This paper is based on the major feature update performed in small batches. After each iteration, the average value of the depth feature is used to determine the center feature [15, 16]. In addition, in order to avoid the fluctuation of the central feature due to wrong samples, the scalar $\alpha$ is set to control the learning rate. The total loss function can be obtained by integrating the loss of the center and the Softmax function, which is expressed by the following formula:

$$L = L_S + \lambda L_C = -\sum_{i=1}^{m} \log \frac{e^{W_{yi}x_i+b_i}}{\sum_j e^{W_{yj}x_i+b_j}} + \frac{\lambda}{2} \sum_{i=1}^{m} \left\| x_i - c_{yi} \right\|_2^2,$$

where scalar $\lambda (0<\lambda<1)$, the weight of the two-loss functions in the entire loss function is primarily adjusted by the loss function representing the center. If the center loss is used to supervise the training, the learned features and centers will degenerate to 0. If only the Softmax loss function is used to supervise the training, the features obtained by deep learning will include many intraclass transformations. Therefore, combining the Softmax loss function with major loss is necessary to complete supervision training, especially for face recognition in TV program production [17, 18].

4.2. Stereo Matching Algorithm Based on Adaptive. By determining the difference of each pixel, the stereo matching algorithm is applied. Disparity refers to the distance between two points. The stereo matching algorithm can be constructed by adaptive cross-window support window, Ad Census adaptive fusion, adaptive region filtering, and disparity refinement, which is discussed in detail.

4.2.1. Construction of Adaptive Cross-Support Window. The variable threshold arm length based on distance linearity is used to compute the criterion using the above face recognition findings. Color threshold $\tau$ decreases linearly with the increase of arm length $l_{pq}$. The calculation criterion of color threshold $\tau$ is as follows:

$$\begin{cases} \tau_1 = \frac{\tau_1 - \tau_2}{L_1} l_{pq}, & 1 \leq l_{pq} \leq L_1, \\ \tau_2 = \frac{\tau_1 - \tau_2}{L_2 - L_1} (l_{pq} - L_1), & L_1 \leq l_{pq} \leq L_2, \end{cases}$$

where $L_1$ and $L_2$ represent the threshold of cross-support distance, $\tau_1$ and $\tau_2$ represent the threshold of color intensity, and $l_{pq}$ represents the distance between the pixel $p$ and the adjacent pixel $q$ waiting for matching.

The change curve of the threshold is shown in Figure 4. In Figure 4, $L_1$ can be understood as the maximum allowable range of the set depth discontinuity area, and $L_2$ represents the set maximum arm length. When $l_{pq} < L_1$ is the threshold of color intensity $\tau$, then the linear generation will decrease. When $L_1 \leq l_{pq} \leq L_2$, a relatively strict threshold change limit is implemented. When estimating the arm length, keep in mind that the color intensity difference between point $p$ and point $q$ must be less than $\tau$. The output value between adjacent pixels $(q, q')$ must be on the arm should and less than $\tau$. The cross-support window and the arm length $\{H_p, H_p'; V_p, V_p'\}$ in four directions of pixel $p$ can be calculated using the constraint criterion.

4.2.2. Ad Census Adaptive Fusion. The adaptive function fusion AD and Census measure based on the shortest arm length are comprehensively considered. The adaptive joint measure function of the AD-Census is

$$C(p, d) = \alpha \left( 1 - \exp \left( -\frac{C_{AD}(p, d)}{\lambda_{AD}} \right) \right) + (1-\alpha) \left( 1 - \exp \left( -\frac{C_{Census}(p, d)}{\lambda_{Census}} \right) \right),$$

$$\alpha = 1 - \exp \left( -\frac{\gamma_L}{L_{min} + \varepsilon} \right),$$

where $\lambda_{AD}$ and $\lambda_{Census}$ represent the regularization parameter, $\alpha$ represents the adaptive weight coefficient, $L_{min}$ represents the shortest arm length in four directions, $\gamma_L$ represents the control parameter of the edge region, and $\varepsilon$ indicates the coefficient of correction.
When the pixel point \( p \) is close to the edge of the image, the shortest vertical arm is short, the Ad weight increases, and the Census weight decreases. The AD and Census measures are combined adaptively.

4.2.3. Adaptive Region Filtering. On the cross-support vertical arm \( \{V_{p}^{-} \sim V_{p}^{+}\} \) of the pixel point \( p \), carry out horizontal expansion of the pixel point \( q \), and the horizontal expansion must meet the color threshold conditions of the horizontal expansion of the pixel point \( P \), expressed as

\[
\tau = \begin{cases} 
\tau_1 - \frac{\tau_1 - \tau_2}{L_1} \tau_2, & 1 \leq l_{q,t} \leq L_1, \\
\tau_2 - \frac{L_2 - L_1}{L_2 - L_1} (l_{q,t} - L_1), & L_1 < l_{q,t} \leq L_2.
\end{cases}
\]

The pixel \( p \) adaptive filtering region can thus be found. The shape of the region can be processed to generate a self-adaptive filtering region, taking into account the smoothness of regional filtering.

\[
R \ominus B = (R \ominus B) \oplus B,
\]

where \( R \) represents the adaptive region after, \( B \) represents the elements of the structure, \( \ominus \) represents the open operation, \( \oplus \) represents corrosion operation, and \( \oplus \) represents expansion operation.

The cost of pixel \( p \) matching in the range \( d \in [d_{min}, d_{max}] \) of parallax is

\[
C(p, d) = \frac{1}{N_{p}} \sum_{q \in N_{p}} C(p, d),
\]

where \( N_{p} \) represents the adaptive filtering region of pixel \( p \), and \( N \) represents the number of pixels in the region. Finally, the WTA method chooses the minimal parallax value of matching cost as the initial parallax in the azimuth parallax.

\[
d_p = \arg \min_{d \in [d_{min}, d_{max}]} (C(p, d)).
\]

4.2.4. Disparity Refinement. Comprehensively consider the region \( N_{p} \) adaptive to the pixel point \( p \) in the image, and select the parallax with the highest statistical frequency in the neighborhood as the parallax of the point \( p \).

\[
d_p^* = \arg \min_{d \in [d_{min}, d_{max}]} (\text{hist}(d, N_{p})),
\]

where \( \text{hist}(d, N_{p}) \) represents the frequency of parallax \( d \) statistics in the neighborhood \( N_{p} \) of point \( p \).

Using consistency detection, the outliers in parallax image are detected.

\[
\begin{align*}
&|d_{\text{left}}(p) - d_{\text{right}}(p - d_{\text{left}}(p))| \geq 1, \\
&d(p) = \min\{d(q_0), \ldots d(q_i), \ldots d(q_{7})\},
\end{align*}
\]

where \( d_{\text{left}}(p) \) represents the visual difference of point \( p \) in the parallax map before optimization, and

\[
d_{\text{right}}(p - d_{\text{left}}(p)) \text{ is the visual difference of the corresponding point of point } p \text{ in the optimized parallax map [19–21].}
\]

The search for outliers can be done in the eight-pixel directions, the minimum parallax pixel is selected, and the parallax image is obtained using median filtering. The implementation of virtual reality technology in TV show production is now complete.

5. Analysis of Experimental Results

To determine the efficiency of using virtual reality technology in TV program production, this research compares it to other methods and carries out simulation experiments. MATLAB is used as the simulation platform. Table 1 shows the specific parameter data of the experiment.

The face recognition methods proposed in this paper are compared with those based on sparse representation and the recognition time of face recognition methods based on deep learning. Figure 5 compares the face recognition accuracy of face recognition methods based on the proposed method to those based on sparse representation and deep learning. Face recognition in AR and ORL face databases is represented in Tables 2 and 3.

We can see that the face recognition method based on YOLOv3 presented in this research has good robustness by...
looking at Tables 2 and 3, and Figure 5. The speed of face recognition proposed in this research is slower in ORL face databases than in AR face databases, as shown in Tables 2 and 3. However, face recognition time in AR face databases and ORL face databases is better than those using sparse representation and deep learning. In Figure 5, the face recognition method based on sparse representation has higher accuracy in the initial recognition. However, when the recognition time increases, the accuracy of facial recognition reduces dramatically. Based on the deep learning face recognition method, the recognition accuracy of the sparse representation-based face recognition method is higher but lower than the YOLOv3-based face recognition method proposed in this paper. The accuracy of the face recognition method proposed in this paper is close to 100%. It demonstrates that the strategy proposed in this paper can improve facial recognition successfully.

Figure 6 shows the adaptive stereo image matching method proposed in this paper and the matching time compared with the image matching algorithm based on feature transformation.

As demonstrated in Figure 6, the adaptive stereo matching technique developed in this research matches substantially faster than the feature transformation-based matching approach. The matching time of the matching algorithm based on feature transformation increases as the number of experiments increases. It indicates that the adaptive stereo matching algorithm proposed in this paper has strong stability and a shorter matching time, making it more suitable for television program production.

This paper applies virtual reality technology to TV program production, and the satisfaction of the psychological need of the audience is higher from the beginning. By analyzing Figure 7, it can be seen that different methods meet the psychological needs of audiences in TV production. The method of reference [4] is compared with the proposed method, which is not high from the beginning of the experiment, nor does it increase significantly with the number of experiments. With the increase in the number of experiments, the fulfillment of the psychological requirements of the audience is gradually improved because this method recognizes faces in TV programs and then uses the adaptive stereo matching method to match images, effectively shortening the program production cycle and improving the quality of TV program production and the program competitiveness.

6. Conclusions

Virtual reality may be used effectively in TV program production in many ways. Several of the concepts discussed here can be implemented right now with little or no new work. Others will take some work, but they are still possible. Virtual reality is now used sparingly in television programs, but this is likely to change fast as engineers understand the promise of this new technology and prices continue to fall. Compared to traditional cinema and television, virtual reality film and television offer a new panoramic experience that can entice viewers to engage in more focused VR experiences. The data can be recognized and tracked continually via the sensor of the IoT, allowing users to try immersive experiences in vision, hearing, and touch. In the production of TV programs, the use of virtual reality technology is vital for improving the overall competitiveness of television programs and meeting viewers’ psychological needs. Furthermore, the application of virtual reality technology in the production of TV programs can also improve the visual effect of TV programs in the context of the new era to achieve the reform of the program form. In order to make greater use of virtual reality technology in the future development of the TV industry, the TV industry must increase its efforts to build up a talented workforce, introduce advanced corresponding equipment, and play an important role in virtual reality technology.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.
Conflicts of Interest

The authors declare that they have no conflicts of interest for publication of this paper.

References


