Hindawi

Research Article

Animation Design Based on 3D Visual Communication Technology

Feng Shan [ID] **and Youya Wang**

Wanjiang College, Anhui Normal University, Wuhu 241008, China

Correspondence should be addressed to Feng Shan; 1301220955@pku.edu.cn

The depth synthesis of image texture is neglected in the current image visual communication technology, which leads to the poor visual effect. Therefore, the design method of film and TV animation based on 3D visual communication technology is proposed. Collect film and television animation videos through 3D visual communication content production, server processing, and client processing. Through stitching, projection mapping, and animation video image frame texture synthesis, 3D vision conveys animation video image projection. In order to ensure the continuous variation of scaling factors between adjacent triangles of animation and video images, the scaling factor field is constructed. Deep learning is used to extract the deep features and to reconstruct the multiframe animated and animated video images based on visual communication. Based on this, the frame feature of video image under gray projection is identified and extracted, and the animation design based on 3D visual communication technology is completed. Experimental results show that the proposed method can enhance the visual transmission of animation video images significantly and can achieve high-precision reconstruction of video images in a short time.

1. Introduction

3D visual communication technology is immersive, interactive, and conceivable for users to create high-quality visual experience. Video based on 3D visual communication technology needs to go through different processes such as content making, coding compression, network transmission, and terminal display. In recent years, visual communication technology has been widely used in various fields because of its function of fast and real-time measurement of information [1]. With the development of multimedia technology, visual communication technology is widely used in the field of video and animation image management, such as network media operation, animation release, etc. Film and television animation and video images are developing constantly in the field of film and television animation and video images. The core of film and television animation and video is visual communication, and it is difficult to obtain target film and television animation and video images in complex background in real time and accurately. It is a common problem in multiframe film and television animation and video visual communication [2].

Reference [3] proposes an innovative design of visual communication system for animated figures and images in virtual reality environment. In the hardware design, the adaptation parameters of the main board of the renderer are designed to optimize the experience of visual communication. In the software design, by introducing Sobel edge operator, the gray function is established to solve the gradient amplitude, and the threshold is selected to compare it to complete the recognition and thinning of the edge data of animated character graphics and images. Design the motion capture module, establish the behavior control model, generate and manage the motion capture files, and complete the overall design of the system. Reference [4] proposes a new optimization method—plane visual communication

effect optimization method based on wavelet change. The image is decomposed by wavelet, and the wavelet is reconstructed. In the process of reconstruction, the modulus diagram and phase angle diagram are calculated, and the edge images of each scale are extracted. The corresponding edge points of the semireconstructed image are enhanced through the edge image. On the above basis, the graphic beautification vector of slip model is used for operation. In order to simplify the operation, the above operation method is transformed into simple mathematical operation, and the visual communication effect is optimized through the reflected light graphic mode. Reference [5] proposes a new automatic graphic language arrangement algorithm in visual communication design. The display size of buffer image in visual communication design is calculated by fixed value method, and the minimum number of layout and the maximum surface utilization are taken as the layout objective function. Ant colony algorithm is used to solve it, get the optimal solution, and obtain the best display position of graphics in visual communication design. The description of graphic language is realized by rule-based syntax description and ASM (abstract state machine) semantic description. The parallel process and selection process of graphic language in visual communication design are described by using parallel and selection marks, and the parallel process and selection process are used to realize the automatic arrangement of graphic language. The proposed algorithm is programmed and run in C language to achieve the purpose of automatic arrangement of graphic language.

However, when the above image visual communication technologies ignore the depth synthesis of image texture, the image visual effect is not ideal. Therefore, a film and television animation design method based on 3D visual communication technology is proposed. Build the video processing process under 3D visual communication technology, and mainly process film and television animated video through 3D video content production, and server and client links; based on the visual distance between 2D image and scene, calculate the zoom factor of 3-dimensional scene transformation; build a deep learning-based deep feature extraction model PCANet, filter images and extract features of film and television animation video images, use PCAN deep network to create a low-resolution feature dictionary, use high-resolution video image reconstruction method, realize LR animated video images and HR animation video images, and introduce nonlocal similarity empirical constraints to optimize HR images; analyze gray and detail features based on wavelet transform, extract the video image features, separate the prospect and background of animated video images, transform K-L (Karhunen–Loeve) as the core, dynamically identify the similarity between evaluation features, and complete the animation design based on three-dimensional visual communication technology. The experimental results show that the proposed method can significantly improve the visual transmission of animated video images and improve the speed of high-precision video image reconstruction.

2. Film and Television Animation Design Based on 3D Visual Communication Technology

3D visual communication technology mainly uses the computer to simulate the environment, to make people feel immersive. Three-dimensional visual communication technology, on the basis of 2D images, reads the main information, compresses and decomposes, and transforms it into three-dimensional images, increasing the measurement and display function of arbitrary sections. In the human-computer interaction module, we can zoom in the three-dimensional image to facilitate users to browse and change the image.

2.1. Video Processing Process under 3D Visual Communication Technology. Figure 1 shows the flowchart of video processing under 3D visual communication technology.

The 3D visual communication technology can be divided into three parts: 3D visual communication content production, server and client, and 3D visual communication content production which is the core of the whole video processing process [6]. In the production of 3D visual communication content, a set of cameras or a set of camera devices and audio sensors containing multiple cameras and sensors are used to capture the sound and visual scene in the real physical world. Cameras and sensors, for example, produce a set of digital video and audio signals. In general, the camera captures content within 360° of the periphery of the unit [7, 8].

Aiming at the video processing under the 3D visual communication technology, the 3D visual communication video image at the same time is combined with mosaic, projection mapping, and frame texture to generate a package, as shown in Figure 2.

2.2. The Scaling Factor of the 3D Scene Transformation. After determining the triangular meshes on a two-dimensional plane through screen projection, a continuous scaling field shall be assigned to each triangular facet; that is, a scaling factor shall be assigned to each triangular facet, and the scaling factor between adjacent triangular facets shall be continuously variable. The specific process is as follows [9, 10]:

S_i is used to represent the distance from the triangle to the viewpoint in 3D scene. S_i is defined as visual distance, and all S_i can be integrated into a threshold of discrete function S_i.

Determine the apparent distance S_i and S_{i+1} between any adjacent triangles i and $i + 1$, and the difference ∂_i between the two can be expressed as follows:

$$\partial_i = \left| S_i - S_{i+1} \right|. \tag{1}$$

Obtain the upper limit value $\max \partial$ of ∂_i through formula (1).

When the value of $\min$ Scale is 0.4, the sample size changes gradually from 0.4–1 to 0.4. Based on the value of $\min$ Scale, pyramid shrinkage is applied to a given texture
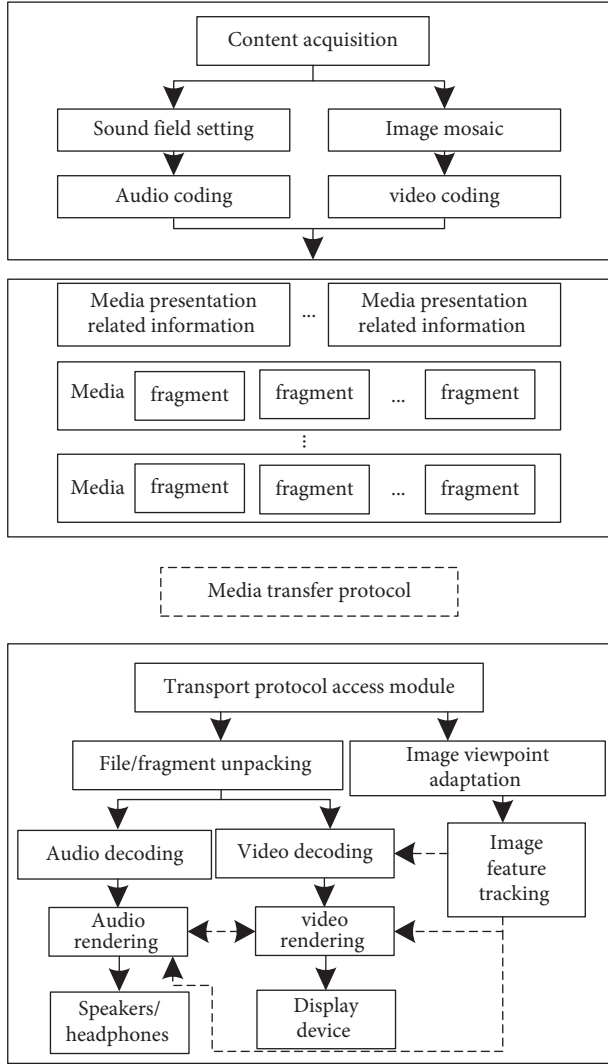
Figure 1: Video processing flowchart under 3D visual communication technology.

Figure 2: 3D visual communication animation video image mosaic, projection, and area-based packaging.

2.3. Deep Feature Extraction Model Based on Deep Learning.
PCANet is a relatively simple method of deep learning based on the theory of deep learning and convolutional neural networks [14, 15]. The PCANet structure is composed of two PCA (principal component analysis) filter layers, a hash layer and a histogram calculation layer, which can extract the deep-seated features of high-resolution animation and video images. But unlike the usual networks, PCANet's filters are more computationally efficient. PCANet is not obtained through training, but through obtaining the mapping results of parts of high-resolution film and television animation video images, and then adopts the PCA (principal component analysis) method to extract the principal components of the high-resolution film and television animation video images, each of which is an independent filter [16]. When the neural network of the filter is used to extract the features, the optimal weights can be obtained without iterative operation; thus, the calculation time is reduced [17]. In order to further discuss the advantages of PCANet depth network in extracting the features of high-resolution animated cartoon video images, a mapping eigenvalue of N, 8×8 and 1 moving distance is set up for the input of animated video images [18, 19]. The main parts of the mapping features L_1 were computed using the PCANet algorithm, and a two-layer filter filters L_1. The steps are as follows:

2.3.1. First Layer: PCA Filter Layer 1.
For the N training picture L_i^N with size of $m1m2$, PCANet selects a $k1 \times k2$ window, and each training image is compared by row and column to obtain the local features of the image. Get $m1 \times m2$ picture blocks of $k1 \times k2$ size. Dislodge the mean value of the picture block, the output behind dislodging the mean value of L_i:

$$\overline{A}_i = \left[\overline{a}_{i,1}, \overline{a}_{i,2}, \cdots, \overline{a}_{i,m_1,m_2} \right]. \tag{2}$$

In formula, each column of the matrix $\overline{A}_i$ represents a block of image, and each column contains $k1 \times k2$ elements. There are $m1 \times m2$ columns in total. After performing the above operations separately for all N training images, a new

sample drawing to obtain a multilevel progressive shrinkage sample drawing, which is used as a scaling factor in the process of texture synthesis to obtain the basis for the corresponding sample drawing [11, 12].

In order to make the neighboring triangle mesh correspond to the neighboring layer, based on max ∂, set up: the value of S function can only be different one layer within the sample layer if the difference is max ∂ that is, according to max ∂ to divide the value range of S function, the value of the function is divided into different layers according to max ∂, and the layers corresponding to the neighboring triangle mesh are also adjacent, thus ensuring the continuity of texture in the process of composition.

The value of S_i for each triangle can obtain a value of Scale (min Scale ≤ Scale ≤ 1) according to the above setting, which is the value of Scale corresponding to the scaling of the meta-texture sample image [13]. Based on this, each Scalevalue corresponds to a Scale-fold shrunk layer, and a Scale-fold shrunk image can be mapped to a Scale-fold shrunk film and TV animation image, where each triangle has a scaling factor.
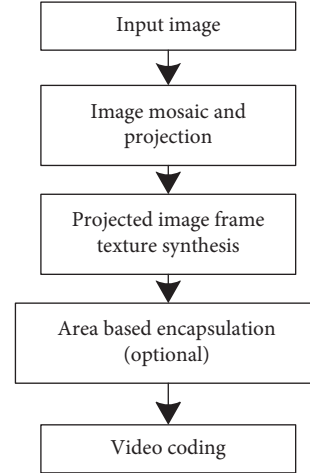
data matrix A was obtained: $A = [\overline{A}_1, \overline{A}_2, \cdots, \overline{A}_N] \in R^{k1k2 \times L}$, and A has $N \times m1 \times m2$ columns. Then, the A matrix was PCA and the ahead L_i eigenvectors of the eigenvector matrix W were taken as the convolutional kernel of the filter group.

$$W_{L_i} = ma_{k1k2}\left(q_l\left(AA^T\right)\right) \in R^{k1k2}, l = 1, 2, \cdots, L_i. \tag{3}$$

In formula, $q_l(AA^T)$ represents the l-th eigenvector of the matrix AA^T. ma_{k1k2} is a function, representing rearranging each of the L_i eigenvectors into the matrix. Each eigenvector contains $k1 \times k2$ elements, amount to get L_i convolutional nuclei of length $k1$ and width $k2$. Finally, for each input image L, the L_i convolutional nuclei were filtered separately. Thus, the first layer of the PCA filter group is completed.

2.3.2. Second Layer: PCA Filter Layer 2.

The second layer of the PCA filter group is the same principle as the first layer. This yields the convolutional kernel of the second layer:

$$W_{L_i}^2 = ma_{k1k2}\left(q_{l_2}\left(BB^T\right)\right) \in R^{k1k2}, l_2 = 1, 2, \cdots, L_i^2. \tag{4}$$

For each input image L_i^2 in the second layer, the PCA filter group is all complete through the filter $W_{L_i}^2$ output features.

Finally, the L_i and L_i^2 feature mappings are obtained in the same way as CNN. PCANet is characterized by these L_1 mapping features, so it can provide a reliable data for the film and television animation video image processing [20]. PCANet uses learning to get multilayer network filters. After the input video image is filtered through two layers of filters, it can have very high-dimensional data output and can be regarded as the input of video image characteristics of the film and television animation and be used for the reconstruction of high-resolution video image of the film and television animation [21].

Therefore, when PCANet extracts the features of film and television animation video images, it is actually a direct processing of the pixels of the film and television animation video images, the operation stage has joined the process of blocking, and the number of film and television animation video image block output from the deep network has increased [22]. PCANet uses the deep network learning process to extract the features of the animation video image, and the video image features extracted by PCANet and the features obtained by artificial rules are more abundant in detailed information and more prominent in texture structure, which provides rich prior knowledge for subsequent reconstruction and fills in the details of the animation video image of low-resolution film and television so as to facilitate the reconstruction of the animation video image of super-resolution film and television.

2.4. Image Reconstruction Method of Multiframe Animation Video Based on Visual Communication.

Combining the advantages of PCANet depth network and sparse representation-based reconstruction method, this paper proposes a visual communication-based multiframe animation video image reconstruction method [23, 24]. For the multiframe animation video image reconstruction method, it is assumed that the high-resolution image and the low-resolution image are sparse expressions for their respective dictionaries, and then, the image sample features are obtained by PCANet depth network. D_h represents the dictionary feature of a high-resolution image, and D_1 represents the dictionary feature of a low-resolution image. In the stage of high-resolution image reconstruction, the low-resolution image is processed in the same way, the deep-seated feature is extracted by PCANet, the sparse representation of D_1 above the low-resolution dictionary feature of the high-resolution image is obtained, the coefficients expressed by the sparse D_1 feature of the low-resolution image are directly applied on D_h, and the corresponding high-resolution feature image is obtained [25]. Low-resolution video animation images are used to achieve high-resolution reconstruction. Through the PCANet depth network, we can get better film and TV animation video image features than the nondepth network. The deep feature dictionary can also improve the description ability of the feature dictionary and improve the quality of film and TV animation video image reconstruction significantly [26].

K blurred film and TV animation video image need to be sampled in the training data set, which is adjusted to the same size to get the corresponding low-resolution film and TV animation video image, which is combined into a model sample pair: $T = \{X_h, X_1\}$, $X_h = \{x_h^i\}_{i=1}^k$ represents high-resolution features, and $X_1 = \{x_h^i\}_{i=1}^k$ represents low-resolution features. Compute the matrix into blocks for all samples in the data set, and select a sliding window of size $k_1 \times k_2$, (normally, the pixel square window of the film and television animation video image used is 3, 5, or 7). After feature extraction for all the film and television animation video images is carried out through the aforesaid sliding window, the new data matrix X of column $N \times m \times n$ can be obtained, and each column of the matrix represents a film and television animation video image block with a total of $k_1 \times k_2$ elements.

To obtain the training sample of i high-resolution animation video image, the formula is

$$\overline{X}_{hi} = \left[\overline{x}_{hi,1}, \overline{x}_{hi,2}, \cdots, \overline{x}_{hi,mn}\right]. \tag{5}$$

Features were extracted with the above data matrix X, and the extracted features were regarded as feature samples in SCSR model and substituted into dictionary features of PCANet.

Suppose that T_{hi}^L represents the results of the high-resolution animated video image after sparse coding, the $H(\cdot)$ function is regarded as a unit step function, the results of sparse coding are quantized, the histogram coding is completed, and the deep detail features of the high-resolution animated video image are extracted, that is, [27, 28]

$$F_{hi} = \left[Bhist\left(T_{hi}^1\right), \cdots, Bhist\left(T_{hi}^{L_1}\right)\right]^T \in R^{\left(2^{L1}\right)L_1 B}. \tag{6}$$

Similarly, it is inferred that through the same processing process as the high-resolution film and television animation video image, the results of extracting the deep-seated detail features of the low-resolution film and television animation video image are given, namely,

$$F_{li} = \left[\text{Bhist}\left(T_{hi}^1\right), \cdots, \text{Bhist}\left(T_{hi}^{L_1}\right) \right]^T \in R^{\left(2^{L1}\right)L_1 B}. \quad (7)$$

Here, F_{hi} and F_{li} on behalf of the film and television animation video image feature extraction results; Bhist on behalf of the histogram coding; and B on behalf of the film and television animation video image segmentation block number of samples.

In this paper, the dictionary is trained in the SCSR framework by combining the sparse coding method. The goal is to obtain a set of dictionary pairs D_h and D_1 that can represent complex feature samples. Make K have the same sparse representation on D_h and D_1 for the deep-level detail features F_{hi} and F_{li} generated by film, television, animation, and video images, and F_{hi} and F_{li} have the same description coefficient, that is,

$$\{D_h, a\} = \|F_{hi} - D_h \cdot a\| + \sum_{i=1}^{K} \lambda_i \|\alpha_i\|,$$
$$\{D_1, a\} = \|F_{li} - D_1 \cdot a\| + \sum_{i=1}^{K} \lambda_i \|\alpha_i\|. \quad (8)$$

In the above formula, $a = \{a_i\}_{i=1}^k$ represents sparse matrix and λ_i represents equilibrium coefficient [29]. In order to make the high-resolution film and television animation video image features and low-resolution film and television animation video image features have the same sparse description method as their respective dictionaries, the combined training method is adopted through formula (5) and (6), that is,

$$\{D_h, D_1, \alpha\} = \frac{1}{N}\|F_{hi} - D_h \cdot a\| + \frac{1}{M}\|F_{li} - D_1 \cdot a\| + \sum_{i=1}^{K} \lambda_i \|\alpha_i\|. \quad (9)$$

In the above formula, N and M, respectively, represent the dimension rearrangement of the column vectors by the element values of high- and low-resolution feature blocks of film, television, and animation video images, and $1/N$ and $1/M$ are the cost between D_h and D_1 of balanced formula (9). In order to facilitate subsequent calculation, formula (10) is reconstituted:

$$\{D_h, a\} = \sum_{i=1}^{K} \lambda'_i \|\alpha_i\|, \quad (10)$$

$$D_C = \begin{bmatrix} 1/\sqrt{N} \, D_h \\ 1/\sqrt{M} \, D_1 \end{bmatrix}, X_C = \begin{bmatrix} 1/\sqrt{N} \, F_h \\ 1/\sqrt{M} \, F_1 \end{bmatrix}, \lambda'_I = \left(\frac{1}{N} + \frac{1}{M} \right) \lambda_i. \quad (11)$$

Equation (11) is solved by iteration. Given dictionary D_C, calculate the sparse representation coefficient a_i of each data sample F_C on D_C, obtain matrix $a = \{a_i\}_{i=1}^k$, and finally update dictionary D_C through a.

After obtaining the dictionary pair D_C, the reconstruction of LR film and television animation video image Y and HR film and television animation video image X can be obtained by using the high-resolution film and television

animation video image reconstruction method based on sparse regular model.

In the stage of reconstructing the video image of film and television animation, some noise exists in the video image of film and television animation due to the interference of external environment noise, and block effect and blurred artifact will appear in the reconstructed video image of film and television animation [30]. Considering that the conventional back-projection model cannot guarantee the quality of the reconstructed image, the details of the image can be better preserved by matching the similar blocks between the image and the image without the prior constraints of partial similarity. Relative to the image block effect and blurred artifacts after the reconstruction, the nonlocal similarity empirical constraint is introduced to optimize the HR image based on the global optimization of back-projection. The global and nonglobal constraint models in this paper are as follows [31, 32]:

$$X^* = \text{argmin}_X \|SHX - Y\|_2^2 + c\|X - X_0\|_2^2 + \gamma\|(I - W)X\|_2^2. \quad (12)$$

In the above formula, S represents sampling operation, H represents fuzzy filtering, $\|X - X_0\|_2^2$ represents global constraint term, $\|(I - W)X\|_2^2$ represents nonlocal self-similarity constraint term, I represents identity matrix, c and γ represent normalization parameters, W represents nonlocal weight matrix, and the element formula in W is as follows:

$$W_{ij} = \frac{1}{z(i)} \exp\left(-\frac{\left\| N(x_i) - N(x_j) \right\|_2^2}{h} \right). \quad (13)$$

In the above formula, $N(x_i)$ represents the i-th film and television animation video image block in the film and television animation video image, x_j represents the searched film and television animation video image block similar to $N(x_i)$, h represents the attenuation factor, and $z(i)$ represents the normalized value.

$$z(i) = \sum_j \exp\left(\frac{\left\| N(x_i) - N(x_j) \right\|_2^2}{h} \right). \quad (14)$$

2.5. Dynamic Recognition of Frame Features of Film, Television, and Animation Video Images

2.5.1. Frame Feature Extraction of Video Image under Gray Projection.
The first step of frame feature extraction is to analyze the features of gray level and detail in the video image based on wavelet transform and then to extract the feature of the video image by gray projection [33]. Finally, the problem of frame feature extraction of film and television animation video image evolves into the problem of foreground and background classification of film and television animation video image, and the separation coefficient is determined by the ratio function of the variance of feature distribution of film and television animation video image in the foreground and background region. The process of

extracting frame feature of film and television animation video image is as follows:

The input of preprocessed grayscale animation video images shall be represented by $f_k(x, y)$, the grayscale image shall be decomposed by Mallat algorithm with wavelet, and a grayscale pyramid $L_k(s)$ shall be established. According to the high-frequency component, the decomposed grayscale image shall be built into detail pyramid $O_k(s, d)$, the size shall be represented by $s \in [0, N]$, and the three detail directions of the video image shall be represented by $d \in [\text{HL}, \text{LH}, \text{HH}]$. The establishment of the original video grayscale image shall be completed by the pyramid $L_k(s)$ and detail pyramid $O_k(s, d)$ calculated in the above process, and the multiple features of two detail features can be obtained simultaneously [34].

In a video image conveyed visually, the local contrast of a certain image at different positions within a certain range is determined by the difference between scales, so as to improve the contrast effect of the animation and video images. The formula is as follows:

$$\text{DOG}(x, y, \sigma_c, \sigma_s) = \frac{1}{2\pi\sigma_c^2}\exp\left[-\frac{r^2}{2\sigma_c^2}\right] - \frac{1}{2\pi\sigma_s^2}\exp\left[-\frac{r^2}{2\sigma_s^2}\right], \tag{15}$$

where r represents the feature contrast measure, σ_c represents the range of feature area, and σ_s represents the range of film, television, and animation video image suppression area. $F(c)$ represents the coarse-scale film and television animation video image, which can suppress the features other than the target area [35]. $F(s)$ represents the fine-scale film and television animation video image, which can describe the detail features of the target area, and then the scale difference between $F(c)$ and $F(s)$ is shown in the following equation:

$$F(c, s) = F(c)\Theta F(s). \tag{16}$$

$F(s)$ adjusts the difference value to the film and television animation video image area with the same scale as $F(c)$, which is represented by $L_k(c)$ and $L_k(s)$, and subtracts each point to obtain the corresponding absolute values $O_k(c, d)$ and $O_k(s, d)$, so as to obtain the gray-level features and detail features of the film and television animation video image. The formula is as follows:

$$\begin{cases} \hat{L}_k(c, s) = L_k(c)\Theta L_k(s), \\ \hat{O}_k(c, s, d) = O_k(c, d)\Theta O_k(s, d). \end{cases} \tag{17}$$

$G_k(i, j)$ represents the gray value of the (i, j)-th pixel in the film and television animation video image, $\hat{d}x$ and $\hat{d}y$ represent the evaluation estimator between the film and television animation video images, and the image feature display diagram is calculated according to the translation estimator, as shown in the following equation:

$$\hat{M}_k(i, j) = \left| G_k(i, j) - G_{k-m}(i - \hat{d}y, j - \hat{d}x) \right|. \tag{18}$$

In the formula, $G_{k-m}(\cdot)$ represents the translation difference, the normalized function (No) is used to unify the scale of the image, and then a group of weighting factors α, β, and γ are determined to complete the weighted fusion of the video image through the above method. The formula is shown in the following equation:

$$S_k = \alpha \cdot N(\hat{L}_k) \cdot \beta \cdot N(\hat{O}_k) \cdot \gamma \cdot N(\hat{M}_k), \tag{19}$$

where $N(\hat{L}_k)$ represents grayscale, $N(\hat{O}_k)$ represents detail, and $N(\hat{M}_k)$ represents the normalization function of motion feature image and represents the feature value of film and television animation video image frame extracted at last.

2.5.2. Dynamic Identification. Dynamic recognition methods mainly include global and local recognition. Because the dimension of video image is high, it is difficult to identify the whole video image by global method, and there are noise interference and information occlusion in the image after dynamic recognition. Therefore, in order to improve the efficiency of dynamic recognition, local recognition method is adopted [36]. This method takes *K-L* (Karhunen–Loeve) transformation as the main core and uses the similarity between evaluation features to realize dynamic recognition.

K-L transform is based on the statistical characteristics of the transformation, through the incoming vector orthogonal transform and the outgoing vector to decouple the relevance of the data, which is the best way in the sense of variance.

Suppose, f_1 represents the $N \times M$-dimensional matrix, Γ represents the sample statistic $\Gamma = \{\Gamma_1, \Gamma_2, \Gamma_3 \dots \Gamma_L\}$, $NM \times 1$ represents the dimension vector, and the average value μ of Γ is calculated and obtained from the set of L sample values, as shown in the following equation:

$$\mu = E(\Gamma). \tag{20}$$

Its variance matrix $C(\Gamma)$ is shown in the following equation:

$$C(\Gamma) = E\left\{(\Gamma - \mu)(\Gamma - \mu)^T\right\}. \tag{21}$$

λ_i represents the eigenvalue of the variance matrix, and Φ_i represents the eigenvector of the variance matrix. $L - 1$ eigenvectors and their corresponding eigenvalues are obtained, as shown in (19):

$$C_\Gamma \Phi_i = \lambda_i \Phi_i, 0 \le i \le L - 1. \tag{22}$$

The order of feature vector Φ_i is based on the order of the size of feature values. The first m eigenvectors are determined and then normalized to form a change matrix $\mathfrak{R}$. Then, the transformation matrix is used to transform vector Γ_i into vector $\mathfrak{I}$ in row m through (20):

$$\mathfrak{I} = \Gamma_i \times \mathfrak{R}. \tag{23}$$

Thus, achieve the effect of dimensionality reduction, and effectively complete the film and television animation video image enhancement design.

3. Experimental Results and Analysis

The simulation environment is Pentium m1.60 Ghz CPU and 760 m RAM. The dual-core Intel second-generation core processor OptiPlex 3010 is used as the main frequency, the system is win10 flagship 32 bits, and the simulation software is MATLAB 2020 B. In order to prove the effect of the method in this paper, it is compared with current advanced visual communication methods of animation images (the plane visual communication design method based on graphic beautification technology and the visual communication design method of animated character image in virtual reality environment).

3.1. Comparative Analysis of Visual Communication Image Reconstruction Effects under Different Methods. To make the experimental results more intuitive, two animated video images of 512×512 pixels shall be selected as reference animated video images, as shown in Figures 3 and 4(a), and the simulation experiment shall be carried out with the relevant simulation experimental tools. If we use Gaussian fuzzy model, set the Gaussian filter of 3×3 region invariant, set its sampling factor to 4, add Gaussian noise to all low-resolution images, and satisfy the signal-to-noise ratio of 30 dB. Figures 3(a) and 4(a) are the reference film and television animation video images, Figures 3(b) and 4(b) are the reconstructions of film and television animation video images by using the visual communication design method of virtual reality environment animation character images, Figures 3(c) and 4(c) are the reconstructions of film and television animation video images by using the graphic beautification technology-based graphic visual communication design method, and Figures 3(d) and 4(d) are the reconstructions of this method.

According to Figures 3 and 4, it can be seen that the visual communication design method of VR environment animated character image cannot avoid blurring edges and the reconstruction effect is not good. The method of plane visual communication design based on graphic beautification technology has a large error when matching some pixels, which leads to the lack of rich and clear details of the final reconstructed animation video image, and the reconstructed effect is inferior to the visual communication design method of virtual reality animation image. Compared with the other two methods, the proposed method has obvious advantages in the reconstruction of high-resolution animation video images.

3.2. Comparative Analysis of Recognition Rate. Compared with the visual communication design method based on VR environment animation and graphic landscaping technology, the effectiveness of visual communication video image dynamic recognition is tested. First, according to the actual situation, 50 kinds of license plate Chinese characters are segmented into two kinds of film and TV animation video images, which are divided into 10 sets of fuzzy similar Chinese characters according to the special structure features of Chinese characters, and then recognized. There are

270 fuzzy film and TV animation video images, and the recognition rates of the two methods are shown in Table 1.

As can be seen from Table 1, the dynamic recognition rate of the proposed method is 91%, which is higher than other methods. The dynamic recognition rate of the graphic beautification technology-based graphic visual communication design method is 62.2%, and the dynamic recognition rate of the virtual reality environment animation figure visual communication design method is 67.5% for the Chinese character images with the same resolution and blurring degree of film and television animation video images. This is mainly due to the fact that the graphic visual communication design method based on graphic beautification technology and the virtual reality environment animation figure visual communication design method do not take into account the influence of illumination color on film and television animation video images, so that the feature dimension is reduced. At the same time, in order to prevent overlapping areas, 2×1 block is used as a block at the time of establishing a block, which also reduces the dimension, but when constructing a histogram, the gradient of each cell is divided into 9 parts. According to the experimental analysis of Dalal, if the area is reduced at this time, the feature dimension can be reduced, but the description ability is also reduced, and the computational complexity is increased. The advantage of the proposed method in recognition rate is that it arranges the eigenvalues in an orderly way and determines the eigenvectors of the eigenvalues as Chinese characters. This method not only reduces the feature dimension, but also retains the key information of the animation video image. In the process of feature extraction, it highlights the main details of the animation video image, deepens the description of the feature, and effectively improves the recognition rate.

3.3. Space-Time Performance Test. Taking the visual communication design method of animated character images in virtual reality environment proposed in reference [3] and the plane visual communication design method based on graphic beautification technology proposed in reference [4] as the comparison method, and taking the grassland texture sample synthesis as an example, the temporal and spatial performance of this method and the comparison method are compared. The results are shown in Tables 2 and 3.

Analysis of Table 2 shows that the length of time of the three methods used in 3D visual communication is different. The length of time of the suggested method is below 3.8 s, the length of time of the visual communication design method of animated character image in virtual reality environment is up to 14.93 s, and the length of time of the graphic visual communication design method based on graphic beautification technology is up to 20.06 s. The length of time of the suggested method is smaller than that in VR environment. Compared with the two comparison methods, this method has obvious advantages in time. This is because this 3D visual communication animation video images through stitching, projection mapping, and animation video image frame texture synthesis projection, resulting in packaging. There is
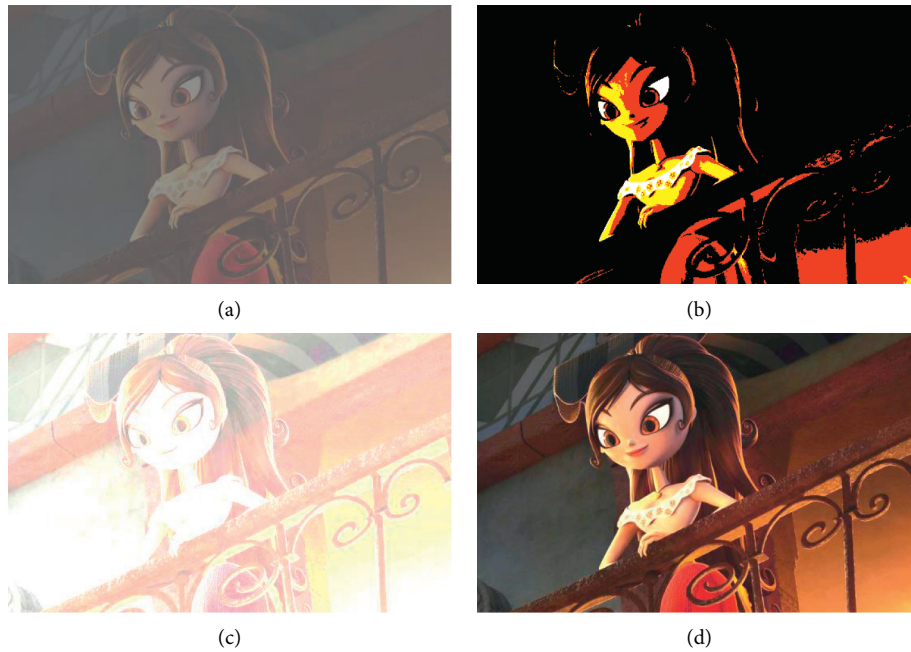
(a)

(b)

(c)

(d)

FIGURE 3: Simulation results. (a) Video image sample of film and television animation. (b) Visual communication design method of animated character image in virtual reality environment. (c) Graphic visual communication design method based on graphic beautification technology. (d) Paper method.
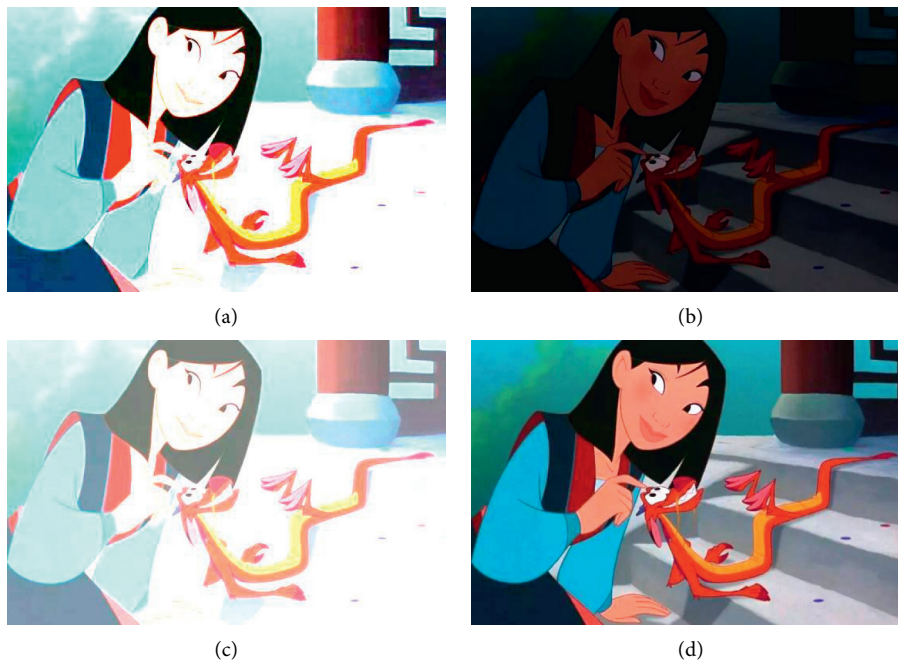

(a)

(b)

(c)

(d)

FIGURE 4: Simulation results. (a) Video image sample of film and television animation. (b) Visual communication design method of animated character image in virtual reality environment. (c) Graphic visual communication design method based on graphic beautification technology. (d) Paper method.

no need to preprocess the texture sample to improve the efficiency of texture synthesis.

Analysis of Table 3 shows that the size of the animation video image after the three methods is quite different. The size of the synthesized image that is determined by this

method is below 3.8 MB, the image size under visual communication design method of animated character image in virtual reality environment is up to 14.93 MB, and the image size under image synthesis based on graphic beautification technology is up to 20.06 MB. The size of the

TABLE 1: Comparison and analysis of recognition rate.

Method	Film animation video image resolution	Characteristic dimension	Recognition rate (%)	Identification time (s)
Visual communication design method of animated character image in virtual reality environment	32×16	70	62.2	1.8
Graphic visual communication design method based on graphic beautification technology	32×16	62	67.5	1.7
Proposed method	32×16	95	91.0	0.9

TABLE 2: Time performance simulation results.

Number of triangular meshes	Synthesis time of this method (s)	Visual communication design method of animated character image in virtual reality environment (s)	Graphic visual communication design method based on graphic beautification technology (s)
512	2.1	16.9	12.8
1024	2.7	22.4	16.7
1536	2.0	23.2	15.0
2048	2.4	34.6	76.2
2560	2.6	51.1	26.6
3072	3.5	87.3	907
3584	3.8	92.2	88.3
4096	3.2	167.2	135.4

TABLE 3: Space performance simulation results.

Number of triangular meshes	The size of the synthesized image is determined by this method (MB)	Image size under visual communication design method of animated character image in virtual reality environment (MB)	Image size after image synthesis based on graphic beautification technology (MB)
512	0.37	1.36	1.78
1024	1.22	2.58	2.81
1536	1.19	4.27	4.89
2048	1.92	5.17	5.74
2560	0.79	7.33	7.63
3072	1.27	9.17	10.08
3584	1.17	11.18	14.27
4096	2.04	14.93	20.06

synthesized image that is determined by the proposed method is smaller than other methods. This is because the present method constructs the scale factor fields for 3-dimensional transformations. Based on this, the identified and extracted video image frame feature scale is smaller and occupies less space.

3.4. Real Example Experiment of Video Image Reconstruction.
In order to reflect the performance of this method in practice, the video image reconstruction performance is tested. The 500 images from the Lion King animated video were selected as subjects and iterated 200 times to test the peak signal-to-noise ratio of the HD reconstructed image, and the results are shown in Figure 5:

As shown in Figure 5, the image peak signal-to-noise ratio of the present method is 31.2dB–40.9 dB, which is higher than the original image, proving that this method meets the practical needs of high-resolution image reconstruction.
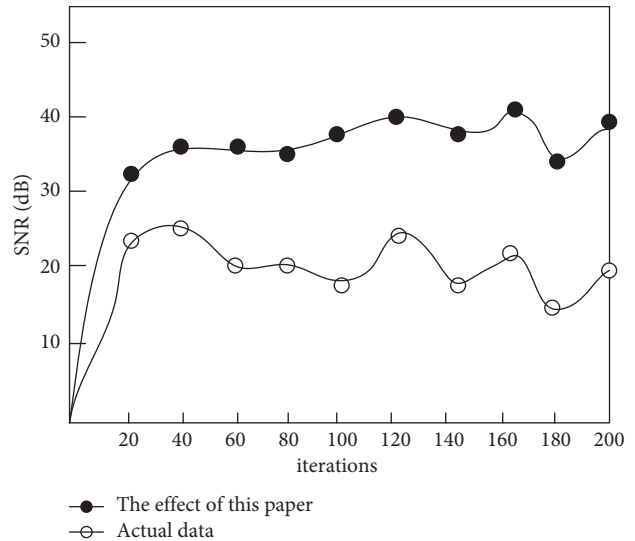

FIGURE 5: Peak SNR of the HD reconstructed images.

4. Conclusion

A new animation design method based on 3D visual communication technology is designed to solve the problem of poor visual effect in the existing image visual communication technology. Build a video processing process under 3D visual communication technology; based on the visual distance from a 2D images to a 3D scene, compute the scaling factor for the 3D scene transformation; building a deep learning-based PCANet model, images were used to filter for processing by using the model and to extract the characteristics of the film and television animation video images; deep features of the image samples were acquired using the PCAN deep network; create a low-resolution feature dictionary for high-resolution images; using a high-resolution video image reconstruction method based on a sparse rule model, realize the reconstruction of animated video images; nonlocal similarity empirical constraints are also introduced to optimize the HR images; analyze gray-scale and detail features based on wavelet transform; extract the features of the video images; with the K-L (Karhunen–Loeve) transformation as the core, the similarity between the picture features is used to dynamically identify the film and television animation video images, and to complete the high-precision film and television animation design. The experimental results show that the proposed method can improve the visual communication effect of animation and video images, with small time cost, small space capacity, and good application performance.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

References

[1] J. Park and B.-U. Lee, "Color image enhancement with high saturation using piecewise linear gamut mapping," *Journal of Visual Communication and Image Representation*, vol. 67, p. 102759, 2020.

[2] M. Yu, "Film and television culture dissemination based on ZYNQ embedded digital image processing," *Microprocessors and Microsystems*, vol. 82, no. 9, p. 103921, 2021.

[3] X. Q. Zhu, "Innovative design of graphic image visual communication system for animated character in virtual reality environment," *Modern Electronics Technique*, vol. 43, no. 10, p. 3, 2020.

[4] M. M. Liu, "Optimization of plane visual communication effect simulation based on graphic beautification technology," *Computer Simulation*, vol. 36, no. 09, pp. 426–429, 2019.

[5] F. Sun, "Automatic programming algorithm of graphic language in visual communication design," *Science Technology and Engineering*, vol. 19, no. 02, pp. 149–154, 2019.

[6] L. U. Mei, Z. D. Wang, and P. Z. Zhang, "A multi-object tracking algorithm based on the combination of neighborhood frame matching and kalman filtering," *Science Technology and Engineering*, vol. 19, no. 15, pp. 204–210, 2019.

[7] C. Rodrigues, Z. Peixoto, and F. Ferreira, "Ultrasound image denoising using wavelet thresholding methods in association with the bilateral filter," *IEEE Latin America Transactions*, vol. 17, no. 11, pp. 1800–1807, 2019.

[8] W. Sun, D. Yan, J. Huang, and C. Sun, "Small-scale moving target detection in aerial image by deep inverse reinforcement learning," *Soft Computing*, vol. 24, no. 11, pp. 1–12, 2019.

[9] I. J. Kadhim, P. Premaratne, and P. J. Vial, "Improved image steganography based on super-pixel and coefficient-plane-selection," *Signal Processing*, vol. 171, pp. 107481–107489, 2020.

[10] M. Meem, A. Majumder, and R. Menon, "Multi-plane, multi-band image projection via broadband diffractive optics," *Applied Optics*, vol. 59, no. 1, p. 38, 2020.

[11] H. Zhang and J. Han, "Mathematical models for information classification and recognition of multi-target optical remote sensing images," *Open Physics*, vol. 18, no. 1, pp. 951–960, 2020.

[12] A. Yuan, X. Li, and X. Lu, "3G structure for image caption generation," *Neurocomputing*, vol. 330, pp. 17–28, 2019.

[13] A. Mik and J. Novák, "Three-component zoom lens with fixed position of optical center," *Applied Optics*, vol. 58, no. 14, p. 3844, 2019.

[14] G. Kertész, S. Szénási, and Z. Vámossy, "Comparative analysis of image projection-based descriptors in Siamese neural networks," *Advances in Engineering Software*, vol. 154, no. 6, p. 102963, 2021.

[15] Y. Feng, H. Wang, Z. G. Qu et al., "Trajectory simulation of hypersonic gliding vehicle based on reference ellipsoid," *Computer Simulation*, vol. 37, no. 10, pp. 18–23, 2020.

[16] K. G. Dhal, A. Das, S. Ray, and J. Gálvez, "Randomly Attracted Rough Firefly Algorithm for histogram based fuzzy image clustering," *Knowledge-Based Systems*, vol. 216, no. 1, p. 106814, 2021.

[17] W. Zhou, Y. Zhou, W. Qiu, T. Luo, and Z. Zhai, "Perceived quality measurement of stereoscopic 3D images based on sparse representation and binocular combination," *Digital Signal Processing*, vol. 93, pp. 128–137, 2019.

[18] J. Yang, W. Zhao, Y. Han et al., "The aircraft tracking based on fully conventional network and kalman filter," *IET Image Processing*, vol. 13, no. 8, pp. 1259–1265, 2019.

[19] N. R. Zhou, X. X. Liu, Y. L. Chen, and Ni-S. Du, "Quantum K-Nearest-Neighbor image classification algorithm based on K-L transform," *International Journal of Theoretical Physics*, vol. 60, pp. 1209–1224, 2021.

[20] S. Gu, L. Wang, W. Hao, Y. Du, J. Wang, and W. Zhang, "Online video object segmentation via low-rank sparse representation," *IET Computer Vision*, vol. 13, no. 5, pp. 469–479, 2019.

[21] A. Rahimi, P. Moallem, K. Shahtalebi, and M. Momeni, "Using Kalman filter in the frequency domain for multi-frame scalable super resolution," *Signal Processing*, vol. 155, pp. 108–129, 2019.

[22] Y. Pu, X. Zhao, G. Chi et al., "Design and implementation of a parallel geographically weighted k -nearest neighbor classifier," *Computers & Geosciences*, vol. 127, pp. 111–122, 2019.

[23] J. Zhao, "Art visual image transmission method based on cartesian genetic programming," *Scientific Programming*, vol. 30, no. 10, pp. 1–10, 2021.

[24] B. Pandey, S. Thakur, H. Joshi, A. Pradhanga, Y. Akiyama, and J. Peethambaran, "Towards video based collective MotionAnalysis through shape tracking andMatching," *Electronics Letters*, vol. 56, no. 17, pp. 2157–2168, 2020.

[25] A. Mustafa, H. Kim, and A. Hilton, "MSFD: multi-scale segmentation-based feature detection for wide-baseline scene reconstruction," *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1118–1132, 2019.

[26] M. K. Sadabad, A. J. Akbarfam, and B. Shiri, "A numerical study of eigenvalues and eigenfunctions of fractional sturm-liouville problems via laplace transform," *Indian Journal of Pure and Applied Mathematics*, vol. 51, no. 3, pp. 857–868, 2020.

[27] H. Ullah, M. A. Khan, and J. Peari, "New bounds for soft margin estimator via concavity of Gaussian weighting function," *Advances in Difference Equations*, vol. 2020, no. 1, pp. 1–10, 2020.

[28] S. Kumar and R. Mathusoothana, "Robust multi-view videos face recognition based on particle filter with immune genetic algorithm," *IET Image Processing*, vol. 13, no. 4, pp. 600–606, 2019.

[29] M. Lee, T. Budavari, I. Sullivan, and A. Connolly, "Sub-band image reconstruction using differential chromatic refraction," *The Astronomical Journal*, vol. 157, no. 5, p. 182, 2019.

[30] Y. Guo, "Moving target localization in sports image sequence based on optimized particle filter hybrid tracking algorithm," *Complexity*, vol. 2021, no. 7, pp. 1–11, 2021.

[31] R. Ma, H. Hu, S. Xing, and Z. Li, "Efficient and fast real-world noisy image denoising by combining pyramid neural network and two-pathway unscented kalman filter," *IEEE Transactions on Image Processing*, vol. 29, no. 3, pp. 3927–3940, 2020.

[32] N. Ghosh, R. Paul, S. Maity, K. Maity, and S. Saha, "Fault Matters: sensor data fusion for detection of faults using Dempster-Shafer theory of evidence in IoT-based applications," *Expert Systems with Applications*, vol. 162, no. 4, p. 113887, 2020.

[33] Z. Ji, Y. Zhang, Y. Pang, X. Li, and J. Pan, "Multi-video summarization with query-dependent weighted archetypal analysis," *Neurocomputing*, vol. 332, pp. 406–416, 2019.

[34] X. Hua, X. Q. Wang, Z. Y. Ma et al., "Video multi-target detection technology based on recursive neural network," *Application Research of Computers*, vol. 37, no. 02, pp. 301–306, 2020.

[35] A. Nicolson and K. K. Paliwal, "Spectral distortion level resulting in a just-noticeable difference between an a priori signal-to-noise ratio estimate and its instantaneous case," *Journal of the Acoustical Society of America*, vol. 148, no. 4, pp. 1879–1889, 2020.

[36] G. Lei, R. Yao, Y. Zhao, and Y. Zheng, "Detection and modeling of unstructured roads in forest areas based on visual-2D lidar data fusion," *Forests*, vol. 12, no. 7, p. 820, 2021.