A Hierarchical Smoothing Method for Animation Image Based on Scale Decomposition

Jieling Jiang(✉) and Wei Li

Guangzhou Academy of Fine Arts, Guangzhou 510006, China
icu007@qq.com

Abstract. In order to solve the problem of low quality of animated images affected by noise, a hierarchical smoothing method for animated images based on scale decomposition is proposed. Get the animation base image, and obtain the detail layer of the source image and target image. Use U-net convolutional neural network to select the decomposition box, select the results according to the remote sensing image segmentation box, and design the image decomposition process. Adjust the animation decomposition scale, focus on measuring multi-scale morphology, use the mean coordinate method to fuse the brightness of the target image, and retain rich details of the image. The fusion image smooth mosaic processing flow is designed, and the minimum variance standard is used to obtain the best matching combination. The gradient is used to represent the direction and size of the pixel changes in the animation image, and the details of the animation image are enhanced by means of superposition correction to achieve image edge smoothing. The experimental results show that the image details obtained by this method are consistent with the image samples, the signal-to-noise ratio is above 90 dB, and the longest smoothing processing time is 27 s, which can obtain high-quality animation images.

Keywords: Scale Decomposition · Animated Image · Hierarchy · Smoothing

1 Introduction

With the rapid development of China's animation industry in the context of modernization, the overall creation form has also undergone tremendous changes, and the corresponding image production method has also been upgraded [1, 2]. At present, the animation industry is developing rapidly, and the use of computer applications for animation creation has also become the main form of animation creation at this stage. Animation is the creation of animation by decomposing the actions of people or objects into many pictures, and then combining these scattered pictures in a certain way to give people a sense of continuous change in vision. The two-dimensional animation is to improve and innovate the traditional animation. Compared with the traditional image design form, the multi frame two-dimensional animation image design has higher flexibility and structural

L. Yun et al. (Eds.): ADHIP 2023, LNICST 549, pp. 20–34, 2024.
https://doi.org/10.1007/978-3-031-50549-2_2

integrity, and is more targeted to the details processing, which to some extent provides great convenience for the subsequent image processing, maintenance and adjustment work [3–5].

In the process of animation image processing, multi scale detail enhancement methods are an effective way to improve the quality of animation. The research on image detail enhancement at home and abroad mainly focuses on enhancing the resolution of animation images in order to improve the design quality of animation images [6]. For this reason, reference [7] proposes a detail blur enhancement method for 3D animation images based on optical parametric amplification. This method first constructs the transmission Relational model of 3D animation image detail features, and then carries out the detail color attenuation processing of 3D animation image through the optical parametric amplification method. The details of the input image are filtered by the filter holding method, the Sobel operator is used to detect the weak edge information of the animation image, and the adaptive interpolation method is used in the edge area to realize the details enhancement of the 3D animation image and the optical parameter detection amplification, and improve the image's fuzzy enhancement and identification ability; A multi-level encoding and decoding image description model based on attention mechanism was proposed in reference [8]. This model uses Faster R-CNN to extract image features, and then uses Transformer to extract three high-level features of the image. The features are effectively fused using a pyramid shaped fusion method. Finally, three long-term and short-term memory (LSTM) networks are constructed to hierarchically decode the features at different levels. In the decoding part, the soft attention mechanism is used to enhance the model to process image edges; Reference [9] proposed an image reconstruction method based on Gaussian smooth Compressed sensing fractional full variation algorithm. This method not only processes the low-frequency components of the fractional order differential loss image, but also increases the high-frequency components of the image, achieving the goal of enhancing image details. The Gaussian smoothing filter operator updates the Lagrange gradient operator to filter out the increase of the high-frequency component of the additive white Gaussian noise caused by the differential operator.

At present, although the traditional methods mentioned above can enhance the details of animated images, they are highly susceptible to external and environmental factors, resulting in unsatisfactory smoothing effects. Therefore, this study proposes a hierarchical smoothing method for animated images based on scale decomposition. This article conducts research from two perspectives: hierarchical scale decomposition and hierarchical smoothing processing of animated images. Based on the results of hierarchical scale decomposition, the image smoothing processing effect and efficiency are improved through scale adjustment, hierarchical fusion, morphological focusing, mean fusion, and smooth stitching processing.

2 Hierarchical Scale Decomposition of Animation Image

2.1 Animation Base Image Extraction

The base signal of the image is called the base image, and the process of preserving the edge image decomposition is the process of finding the base image [10].

For the extraction of the base image, the problem of preserving edges in the base image is considered first. For one-dimensional signals, the existence of edges can generally be determined by the four adjacent extreme points [11, 12].

The workflow on the brightness layer is shown in Fig. 1.

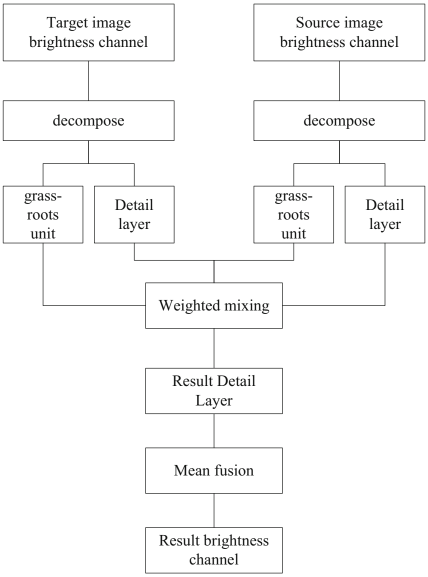


Fig. 1. Workflow on target image decomposition brightness layer

First, the source image and target image are converted to CIELAB color space, and then the brightness channel is decomposed into base layer and detail layer. The base layer is a structural layer containing image structure information. Edge smoothing methods such as bilateral filtering and weighted least squares filtering can realize the decomposition problem of the base layer and detail layer [13, 14]. For this reason, the weighted least squares method is used, because it has excellent performance in solving the blur level near the edge. Given a brightness channel, the base layer in the channel can be obtained by solving the following minimum values:

$$\min S = \sum_{m} (H(m) - O(m))^2 + \alpha \left(\frac{|H_x(m)|^2}{|O_x(m)|^\lambda} + \frac{\left|H_y(m)\right|^2}{\left|O_y(m)\right|^\lambda} \right) \tag{1}$$

In formula (1), m represents each pixel in the fusion area, H represents the smoothed base layer, O indicates the image brightness channel, $O_x(m)$, $H_x(m)$ means on x gradient in direction, $O_y(m)$, $H_y(m)$ represents the gradient in the Y direction [5]. The first is to make H try to connect with O keep similar, the second item is to minimize H yo smooth it. coefficient α it is used to balance the two: α the higher the value of, the smoother the base layer will be subscript x, y the partial derivatives are calculated on the abscissa and ordinate respectively coefficient λ decided O sensitivity of gradient.

Based on this, the algorithm uses subtraction to obtain the detail layers of the source image and the target image respectively:

$$C = O - H \tag{2}$$

In formula (2), C represents the detail layer of the image. The brightness balance can be achieved by transferring the brightness histogram of the target image base layer to the source image base layer, and the hole repair method is used to modify the detail layer of the target image in the fusion area [15, 16].

2.2 Image Hierarchical Scale Decomposition

Based on the obtained animation base image, the U-net convolutional neural network selection method is used in the image scale decomposition process, but the size of the decomposition box is not suitable for image decomposition. If only the original anchor box size is used for learning, the decomposition block will be misjudged, and its average identification rate is high, so the anchor box size needs to be corrected. By optimizing the given decomposition framework, the decomposition framework can cover the target in the image well [17, 18].

Determine the length, width, height and rotation angle of the center point of the bounding box according to the calculated position of the animation base image decomposition box, that is $box\{l, c, h, \theta\}$ after adding the rotation angle to the original decomposition box, calculate the decomposition box information, which is expressed as:

$$\begin{cases} \gamma_l = \Delta l \cdot \iota_l + \iota_l \\ \gamma_c = \Delta c \cdot \iota_c + \iota_c \\ \gamma_h = \Delta h \cdot \iota_h + \iota_h \\ \gamma_\theta = \Delta \theta \cdot \pi + \iota_\theta \end{cases} \tag{3}$$

In formula (3), $\Delta l,\ \Delta c,\ \Delta h,\ \Delta \theta$ respectively represent the displacement between the four vectors and their upper left positions; $\iota_l,\ \iota_c,\ \iota_h,\ \iota_\theta$ represent the minimum decomposition frame information of the surrounding target of the above four vectors respectively. Based on this, the candidate deviation value of the decomposition box is obtained, and the decomposition box is precisely determined according to this value [19].

Select the results according to the remote sensing image segmentation box, and design the image decomposition process:

In the decomposition area, the distance between the decomposed image blocks is represented by the color variance information:

$$d = \sum_{n=1}^{m} \left(\varepsilon_{n,m} - \left(u_1 \sigma_{n,m} + u_2 \varepsilon_{n,m} \right) \right) \tag{4}$$

In formula (4), n represents the set of adjacent points in the decomposed region image; $\varepsilon_{n,m}$ represents the mean square deviation of each parameter of the image; u_1,

u_2 represent the number of pixels of two images respectively [20, 21]. Using this equation, it can ensure that after decomposition, regions with small color variance can be decomposed first, while maintaining the same characteristics of colors, and the histogram characteristics of regions can be reflected in a sense.

3 Hierarchical Image Smoothing Based on Scale Decomposition

3.1 Animation Decomposition Scale Adjustment

According to the hierarchical fusion results of the scale decomposition image, the basic enhancement nodes are laid out by integrating the image detail enhancement requirements and standards, and corresponding associations are formed within a reasonable range.

First, use the test animation image of professional device and equipment atlas, and use the basic color of details to replace the dark primary color to set the image gray scale unchanged. Secondly, calculate the relative value of pixel depth of field, and set the image detail enhancement spacing as the basis. Finally, the relevant enhancement nodes are set according to the change of the initial pixel transmission intensity. When setting nodes, it is necessary to ensure that nodes are interrelated and can form a circular detail enhancement program [22, 23]. The recognition efficiency of the image can be further enhanced after the overlapping correlation is made for the change of the frame number of the animation image and the positioning of the enhancement details. Set marks at the location of enhancement defects to complete initial directional enhancement inspection and enhancement node setting and control.

The image recognition degree and frame number of animation are fixed, and animation basic decomposition is required for changes in requirements and standards. First, the intercepted animation is described and identified in a basic way, then the animation is separated, and finally the separation step is calculated. The specific formula is:

$$\phi = (B + E)^2 \times \frac{\delta \times \sum_{i=1} (e_i + 1)}{B} \tag{5}$$

In formula (5), ϕ is the animation separation step size; B is the preset total range; E is the stacking range; δ Is the resolution deviation. According to formula (5), the animation separation step size can be measured and set as the basic animation classification standard. The difference between the separation step size of two orthogonal processing directions and the scale of the original image is 1/4. After the design and adjustment of the principle of identifiable animation features are completed, the preset animation decomposition scale is constantly adjusted according to the needs, which can create a stable environment for the details enhancement of subsequent images.

3.2 Scale Decomposition Image Hierarchical Fusion

Due to the limited depth of field of the camera's optical lens, the optical lens can only capture images focused on local scenes. Therefore, only objects within the depth of field are focused and clear, while objects outside the depth of field are blurred. However, the information transmitted by partially focused images is incomplete, because not all meaningful objects are focused in one image. When the boundary between the image focus area and the defocus area is complex, image fusion processing needs to be carried out.

This study completes the hierarchical fusion processing of scale decomposed images from the perspectives of multi-scale morphological focusing measurement and parallel mean fusion.

3.2.1 Multi Scale Morphological Focusing Measurement

In image processing, gradient represents the sharpness information of the image, and morphological gradient operator can well extract the gradient information of the image, and can expand to multi-scale morphological gradient operator by changing the size of structural elements, and carry out filtering and other operations. Similarly, the multi-scale morphological gradient operator is used to extract the gradient information of the image at different scales, and then these information gradients are integrated to form an effective focus measurement, which is called multi-scale morphological focus measurement [24]. Build many scale structure elements, which can be expressed as:

$$A_n = a_1 \oplus a_2 \oplus \cdots \oplus a_n \tag{6}$$

In formula (6), $\oplus$ denotes the expansion operator in the morphological gradient operator; $\{a_1, a_2, \cdots, a_n\}$ express n basic elements a a collection of components. In mathematical morphology, structural elements are virtual tools used to extract image features, and structural elements of different shapes are used to extract different types of image features. In addition, changing the size of structural elements can expand to multi-scale, and these multi-scale structural elements can be used to extract comprehensive gradient features in the image.

Using morphological gradient operator to calculate image mesoscale as n the gradient characteristics of are:

$$D_n(x, y) = f(x, y) \oplus A_n - f(x, y) \otimes A_n \tag{7}$$

In formula (7), $f(x, y)$ image; $\otimes$ represents the erosion operator in the morphological gradient operator. In morphological operations, morphological gradient is equal to expansion operator minus erosion operator.

Morphological gradients at different scales are integrated into multi-scale morphological gradients, and multi-scale morphological gradients are constructed by integrating morphological gradients at different scales with weighted sum. Different weighting values are allocated under different scales. The larger the scale, the smaller the weighting value. On the contrary, the smaller the scale, the larger the weighting value. The integrated weighted gradient map can express the gradient information well, and can also clearly and effectively transfer the focus information of the source image.

The multi-scale morphological gradient in the region is summed to construct the multi-scale morphological focus measurement of the region. Since the sum of gradients helps to measure the sharpness of the area and suppress noise, the focus measurement of the area is described by the sum of multi-scale morphological gradients as follows:

$$L_n(x, y) = \sum_{n=1}^{k} \omega_k D_n(x, y) \tag{8}$$

In formula (8), ω_k express k weighted values. In the whole process of multi-scale morphological gradient measurement, it is also necessary to set the shape and scale values of structural elements. The shape of the structural element determines the shape information of the signal extracted by the operation. Different results can be obtained by processing images with different structural elements.

3.2.2 Parallel Mean Fusion

Once the base layer and detail layer information of the source image is obtained, the new source image brightness channel can be calculated from the channel. Then the new brightness channel is fused to the brightness layer of the target image using the mean coordinate method. Based on this, build a hierarchical image fusion model, as shown in Fig. 2.

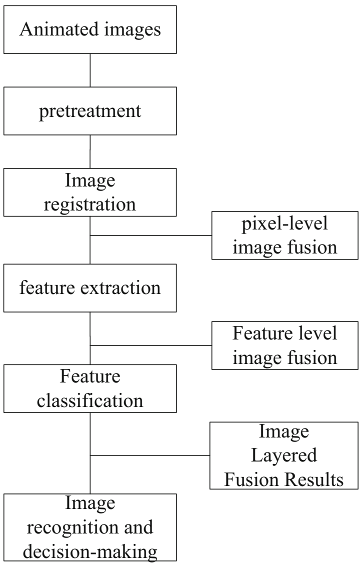


Fig. 2. Hierarchical image fusion model

Given a region to be fused and its chain of boundary pixels given in counterclockwise order, the mean coordinate of each pixel is the weight, and is calculated by the following formula:

$$\omega_i = \frac{\tan(\beta_{i-1}/2) + \tan(\beta_i/2)}{\|v_i - \vartheta\|} \tag{9}$$

In formula (9), β_i indicate the process i angle obtained by the second cycle; ϑ represents pixels. Using the mean coordinates of these pixels, the mean interpolation can be defined as follows:

$$\varsigma(\vartheta) = \sum_{i=1}^{m} \omega_i(\vartheta)\left(O_i - O_i'\right) \tag{10}$$

In formula (10), O_i' represents the new source image brightness channel.

Because the mean interpolation of each pixel in the fusion area is only related to the boundary coordinates, the interpolation of each pixel can be processed separately. It takes time to implement mean interpolation on CPU. In order to solve this performance bottleneck, CUDA technology is used to accelerate interpolation. The results obtained through the new source image brightness channel may exceed the brightness range, which may result in the loss of the visual effect of the image, so these values should be converted to the display range. The algorithm migrates the intensity histogram of the target brightness to the current brightness, and successfully retains the rich details of the image.

3.3 Smooth Mosaic Processing of Fused Images

First, create a matching relationship for the extracted image contour features. The image contour features taken include point features, edge features, contour features and area features. Match the image contour features with the boundary correlation constraint method. After matching the matched contour feature points, many matching pairs will appear near the boundary of the point. Therefore, it is necessary to detect the matching process by introducing the information of contour feature vector. The specific process is shown in Fig. 3.

After extracting and processing the image contour features, the image contours to be spliced have corresponding contour descriptors, and the minimum variance standard is used to obtain the best matching combination.

3.4 Image Edge Smoothing

After completing the histogram equalization of the background layer image, multi-scale animation images are fused by bilateral filtering, which mainly uses the edge stability of the bilateral filtering image and the translation invariance of the direction filtering to capture the feature structure of the image and avoid the loss of edge details or excessive noise after image fusion. The gradient weight factor is introduced to apply the gradient estimation result to the bilateral filtering. The gradient quoted here represents the direction and size of the pixel changes in the animation image.

The multi-scale edge of the image is perpendicular to the gradient direction, and the pixel with a smaller angle between the two included angles will eventually output a smaller weight. The size and direction of the pixel value change in the above are mainly obtained by derivation calculation, which can also reflect the contrast and change trend of image edge pixels. Calculate the horizontal partial derivative array and the vertical partial

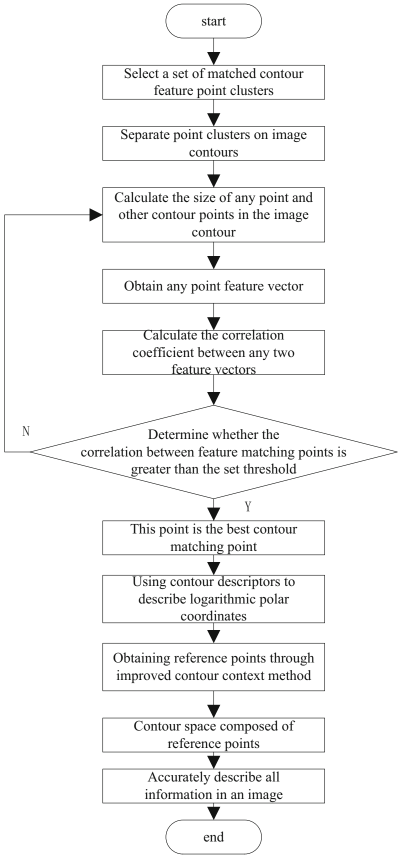


Fig. 3. Fusion image smooth splicing process

derivative array respectively. After obtaining the calculation results, use the second order norm to calculate the pixel gradient size and direction.

After the image gradient calculation is completed, considering the local direction information of the image edge contained in the gradient vector and the information perpendicular to the gradient direction, the gradient function is used to process the image edge pixels. In the process of multi-scale decomposition, the pixels in the vector direction occupy a high weight, so a gradient kernel function is constructed, and the gradient value corresponding to the image edge pixel points is calculated by using the first order partial derivative. The calculated gradient value is substituted into the gradient kernel function, and the pixel angle included in the gradient vector is calculated.

Then, the details of the animated image are enhanced by means of superposition correction. First, the scanning animation image is recognized by light shadow mapping, and the key frame number position and recognition features are marked. Secondly, set the optical parameter correction structure to analyze the change of optical parameters through the change of image spectrum. Finally, in case of any abnormality in the image detail enhancement process, the full coverage positioning shall be carried out immediately, and the correction amount of changes shall be adjusted evenly to ensure the correction effect of the detail enhancement, and the image recognition degree and frame number shall be set. The above process needs to set a directional correction edge value, and the calculation formula is:

$$\mu = \sum_{i=1}^{m} \left[(\upsilon \times 0.5b)^2 - \phi \right] \tag{11}$$

In formula (11), υ represents the pixel transmission intensity; b indicates stack offset. The corrected edge value obtained from the test is taken as the measurement standard of animation image correction processing, and the visual feature recognition model is comprehensively constructed to synchronously enhance the application quality of animation images, so as to achieve image edge smoothing.

4 Experiment

In order to verify the practical application effect of the hierarchical smoothing processing method of animation image based on scale decomposition, and considering the rigor and scientificity of the experiment, the comparative experiment is designed. The experimental methods are the fuzzy enhancement method of 3D animation image details based on optical parametric amplification, the image description model of multi-level encoding and decoding based on attention mechanism. The image reconstruction method based on Gaussian smooth compression perception fractional order total variation algorithm and the method proposed in this paper. In the experiment, the smoothness of image details is taken as the experimental item, and the image signal-to-noise ratio and processing time are taken as the comparison experimental criteria. The performance of the smoothing methods is compared and analyzed from this aspect.

4.1 Experimental Conditions

The image data used in the experiment is mainly from the animated image Standard library ImageNet, which contains animated images of people, landscapes, animals, plants, etc. This is a very large image dataset that contains over one million high-resolution images and covers thousands of categories. The ImageNet dataset is widely used in research on computer vision tasks such as image classification, object detection, and image segmentation. In the experiment, color images with unclear details in various categories are mainly selected. The selection of animation image samples mainly includes the following characteristics. From the aspect of image elements, image samples must include natural scenes or artificial scenes; From the aspect of image illumination, the selection of image samples should consider two situations: sufficient illumination and insufficient illumination; In terms of image color, image samples need to change from color rich scenes to color free scenes.

According to the above analysis, 600 animated images that meet the above requirements are used for testing, including 100 building images, people images, tree images,

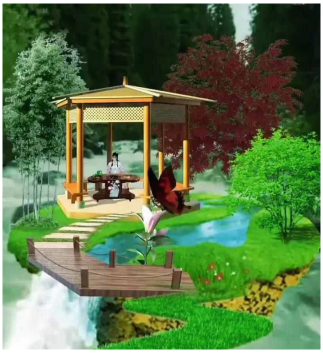


Fig. 4. Image sample

dim light images, single color images, and rich color images. One image sample is shown in Fig. 4.

In the experiment, the i5-3570K computer was used as the platform, and the experimental results were obtained through third-party software.

4.2 Experiment and Analysis of Image Detail Smoothing Performance

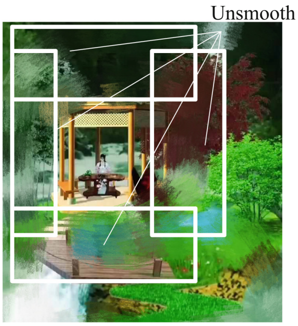
Image detail smoothing is the basic indicator for effective verification of image hierarchical smoothing processing. Therefore, during the experiment, the detail smoothing performance of four different processing methods is compared and analyzed, as shown in Fig. 5.

It can be seen from Fig. 5 that the image detail information obtained by using the studied method is consistent with the image sample, while the details of the other three methods are blurred, indicating that the image details are not smooth.

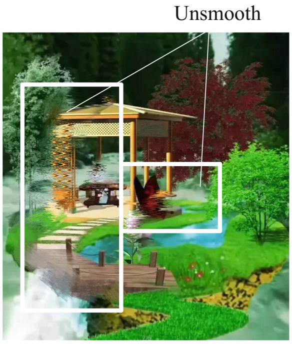
4.3 Experiment and Analysis of Image Signal-to-Noise Ratio

For noisy animation images, the signal-to-noise ratio is a standard evaluation index. The larger the signal-to-noise ratio, the better the image quality is, and the better the hierarchical smoothing effect of the image is. Based on the above experimental results, the signal-to-noise ratio of image samples processed by different methods is calculated to measure the denoising level of different methods. The signal-to-noise ratio calculation results of different methods are shown in Table 1.

The results in Table 1 show that under the experimental conditions of animation images with different calculation times, the signal-to-noise ratio of the three traditional methods is below 75 dB, which indicates that the image contains more noise; In contrast, the signal-to-noise ratio of the designed scale decomposition method is more than 90 dB, indicating that the image quality is better and there is almost no noise interference. According to the experimental results of image detail smoothness, the designed method has smoother detail processing, better denoising effect, and higher quality and better hierarchy of animation images.



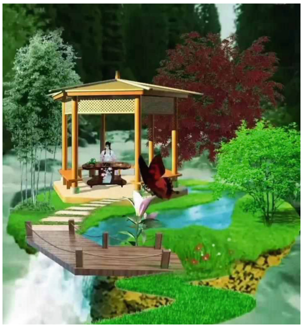
(a) A Method of 3D Animation Image Detail Fuzzy Enhancement Based on Optical Parametric Amplification



(b) An Image Description Model Based on Attention Mechanism and Multilevel Coding and Decoding

(c) Image reconstruction method based on Gaussian smooth compressed sensing fractional order total variation algorithm

Fig. 5. Comparison and Analysis of Image Detail Smoothing Performance of Different Methods



(d) A Hierarchical Smoothing Method for Animation Image Based on Scale Decomposition

Fig. 5. (*continued*)

Table 1. Signal to noise ratio calculation results of different methods/dB

Calculation times /Times	Optical parametric amplification method	Attention mechanism method	Gaussian smooth compressed sensing method	Scale decomposition method
2	70	60	52	91
4	68	58	50	92
6	67	59	48	92
8	68	58	49	91
10	69	57	50	90
12	68	59	51	91
14	70	60	52	92
16	68	55	52	93
18	69	56	53	92
20	68	58	50	93

4.4 Experiment and Analysis of Image Hierarchical Smoothing Processing Time

Compare the image hierarchical smoothing processing time of different methods, and the comparison results are shown in Fig. 6.

It can be seen from Fig. 6 that the longest smoothing processing time of the 3D animation image detail blur enhancement method based on optical parametric amplification, the image description model based on attention mechanism multi-level coding and decoding, and the image reconstruction method based on Gaussian smooth compression perception fractional order full variation algorithm are respectively 140 s, 175 s, 200 s, and the longest smoothing processing time of the studied method is 27 s.

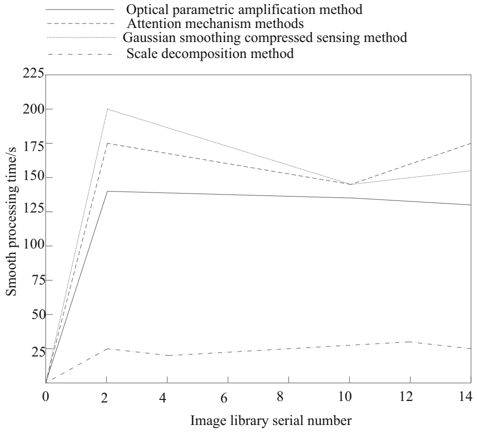


Fig. 6. Hierarchical smoothing processing time of images by different methods

5 Conclusion

A hierarchical smoothing processing method of animation image based on scale decomposition is proposed. This method uses scale decomposition method to produce visually realistic fused image details, integrates highly parallel mean fusion technology, and realizes hierarchical smoothing processing of animation image. In a complex animation processing environment, it can quickly identify the problems existing in details, and further improve the overall smoothing effect. Ensure the detail expression ability of images and create a more stable innovation environment. Experimental results show that the proposed method can match the target image well in detail.

Although this algorithm has the above advantages, it still has some limitations. For example, the current implementation is only partially parallel and region filling consumes a lot of CPU time. Therefore, in the future work, we should try to use GPU to handle the whole process, so as to improve the computational efficiency of the algorithm.

References

1. Gao, C., Song, C., Zhang, Y., et al.: Improving the performance of infrared and visible image fusion based on latent low-rank representation nested with rolling guided image filtering. IEEE Access **16**(5), 1–14 (2021)
2. Chen, H., Deng, L., Qu, Z., et al.: Tensor train decomposition for solving large-scale linear equations. Neurocomputing **464**, 203–217 (2021)
3. Chew, A., Ji, A., Zhang, L.: Large-scale 3D point-cloud semantic segmentation of urban and rural scenes using data volume decomposition coupled with pipeline parallelism. Autom. Constr. **133**(01), 1–19 (2022)
4. Zhang, X., Yan, H.: Medical image fusion and noise suppression with fractional-order total variation and multi-scale decomposition. IET Image Proc. **15**(8), 1688–1701 (2021)
5. Ren, L., Pan, Z., Cao, J., et al.: Infrared and visible image fusion based on edge-preserving guided filter and infrared feature decomposition. Signal Process. **186**(2), 108–116 (2021)

6. Chen, J., Wu, K., Cheng, Z., et al.: A saliency-based multiscale approach for infrared and visible image fusion. Signal Process. **182**(4), 107–118 (2021)
7. Yanxia, Y.: Detail blur enhancement method of 3D animation image based on optical parametric magnification. Laser J. **43**(09), 114–118 (2022)
8. Li, K., Zhang, J.: Multi-layer encoding and decoding model for image captioning based on attention mechanism. J. Comput. Appl. **41**(09), 2504–2509 (2021)
9. Yali, Q., Jicai, M., Hongliang, R., et al.: Image reconstruction based on Gaussian smooth compressed sensing fractional order total variation algorithm. J. Electron. Inf. Technol. **43**(07), 2105–2112 (2021)
10. Madureira, A.: Hybrid localized spectral decomposition for multiscale problems. SIAM J. Numer. Anal. **59**(2), 829–863 (2021)
11. Wang, K., Wang, S., Sun, Q., Liu, C., Chen, S.: Point cloud segmentation matching for 3d reconstruction using multi-layer lidar. J. Changchun Univ. Sci. Technol. **43**(04), 49–56 (2020)
12. Li, Y., Wang, J.: Edge feature extraction method of Brillouin scattering spectral image. Opt. Commun. Technol. **45**(03), 37–41 (2021)
13. He, L., Su, L., Zhou, G., Yuan, P., Lu, B., Yu, J.: Image super-resolution reconstruction based on multi-scale residual aggregation feature network. Laser Optoelectron. Progress **58**(24), 250–259 (2021)
14. Li, C. Research on optimization of 3D image enhancement based on adaptive. Comput. Simul. **37**(12), 358–361 (2020)
15. Liu, W., et al.: Research on intelligent image processing based on deep learning. Autom. Instrument. **12**(08), 60–63 (2020)
16. Zhao, S., Yang, T.: A coherent coefficient based filter of complex number images. Comput. Technol. Develop. **30**(02), 7–11 (2020)
17. Dr. Neetu, A.: Image recognition through human eyes, computers and artificial intelligence. J. Res. Sci. Eng. **3**(3), 132–141 (2021)
18. Duan, H., Wang, Z., Wang, Y.: Two-channel saliency object recognition algorithm based on improved YOLO network. Laser Infrared **50**(11), 1370–1378 (2020)
19. Ajay, R., et al.: Computer vision and machine learning for image recognition: A review of the convolutional neural network (CNN) model. Asian J. Multidimen. Res. **10**(10), 1023–1029 (2021)
20. Jin, H., Cao, T., Xiao, C., Xiao, Z.: Video summary generation based on multi-feature image and visual saliency. J. Beijing Univ. Aeron. Astron. **47**(03), 441–450 (2021)
21. Jeya, C.A., Dhanalakshmi, K.: Content-based image recognition and tagging by deep learning methods. Wireless Pers. Commun. **123**(1), 813–838 (2021)
22. Xiang, J., Xv, H.: Research on image semantic segmentation algorithm based on deep learning. Appl. Res. Comput. **37**(S2), 316–317+3 (2020)
23. Li, M., Li, L., Lei, S.: Application of unsupervised fuzzy clustering algorithm in image recognition. Techn. Autom. Appl. **39**(01), 121–124+159 (2020)
24. Li, P., Li, J., Wu, L., Hu, J.: Image recognition algorithm based on threshold segmentation method and convolutional neural network. J. Jilin Univ.(Science Edition) **58**(06), 1436–1442 (2020)