An Adaptive Texture Enhancement Algorithm for AR Live Screen Based on Approximate Matching

Panpan Yang, Lingfei Ma, Chao Yin*, Yang Ping
Institute of Data Science and Agricultural Economy
Beijing Academy of Agricultural and Forestry Sciences, Beijing, 100097, China

***Abstract*—In order to improve the visual effect of AR live video detail texture, an adaptive enhancement algorithm based on approximate matching is proposed. According to the local self-similarity of the original image, the best matching block of the initial super-resolution image block is obtained. After extracting its high-frequency information, the improved singular value decomposition (SVD) is used to embed the watermark into the original super-resolution gray image; And through the watermark extraction and reconstruction matrix, the effective high-frequency detail texture information is obtained, then the final super-resolution image is obtained, and the AR live video detail texture enhancement is completed. The experimental results show that the peak signal-to-noise ratio, structure similarity and feature similarity of the algorithm are high, and the reconstruction effect is good; After detail texture enhancement, the detail texture clarity and edge sharpness of AR live picture are better, and the best visual effect is achieved.**

***Keywords*—Approximate matching; AR live; picture details; adaptive; texture enhancement; matrix reconstruction**

I. INTRODUCTION

AR, also known as augmented reality, is a fusion technology used between the real and virtual worlds. It can realize the real environment and virtual objects overlap, so that both appear in the same screen. It has 5 core technologies, namely, tracking registration, display, virtual object generation, interaction and merging [1]. AR live broadcast is the product of this technology, which uses cameras and sensors to collect information and data in real environment. After the input processor is analyzed and reconstructed, it is returned to the display terminal such as the camera, And AR live terminal screens can have multiple special effects [2]. AR live images usually contain three categories, such as flower, car screen type images, human-dominated portrait images and scene images that can sense the environment. When these pictures are presented through a display terminal such as a camera, the details and texture of the live pictures will be affected by the network, device itself and other problems [3]. The visual homogeneity in the picture is the detailed texture of the picture, and it also represents a feature of the picture. Texture can give the picture a depth or can be understood as an object of visual presentation. Approximate matching refers to the image which has the highest similarity degree and can be matched with the original image by obtaining the most similar image, and can retain the original image features.

In order to ensure the visual effect of the detailed texture of AR live images, the algorithms based on guiding coefficient weighting, NSST multi-scale and deep learning are proposed respectively in [4], [5] and [6] to improve the visual effect of images. Guide coefficient weighting algorithm in the enhancement process. Mainly through the brightness of the image and non-brightness region to complete the division of the screen processing. NSST multi-scale algorithm is the implementation of low-light picture conversion to complete the screen processing. Deep learning algorithm adopts the strategy of mixing deep learning and image fusion to improve the image quality. In the process of processing, the three algorithms do not extract and reconstruct the detailed texture features, so the clarity of the enhanced texture is poor. Based on this, this paper uses approximate matching algorithm to enhance the detail texture and improve the clarity of detail texture.

II. APPROXIMATE MATCHING TEXTURE ADAPTIVE ENHANCEMENT ALGORITHM

A. Approximate Matching of AR Live Broadcast Screen Detail Texture High Frequency Component Extraction

The pictures in the local neighborhood have similar picture image blocks, which is an obvious feature of the local self-similarity of AR live broadcast pictures. And there are the same scale or different scales in the same picture. Correlation analysis and research are carried out on this feature, and the high frequency detail of similar blocks are extracted and analyzed. The detailed texture of the picture is processed by synthetic method. High frequency detail is obtained from high resolution images, and super- resolution reconstruction is based on local self-similarity. The original live picture I, its corresponding high frequency live picture HF and its corresponding interpolation magnified live picture LR are used as the training set to generate the super- resolution AR live picture HR.

The initial high-resolution image LR is obtained by the contour template interpolation algorithm, and I is amplified after the implementation of processing, and the magnification is α. n order to avoid the missing of the image detail texture, it is necessary to determine the window in LR, which is represented by small window image block e. And formed in the case of pixel (i, j) as the center; If a similar block is retrieved at the same time, belongs to a large window of q, and is in I's position for (x, y), then:

*Corresponding Author.

$$x = i/\alpha, y = j/\alpha \tag{1}$$

In formula: The scaling factor is expressed as α. Through matching window e in the search window f near the scope of similar block search, access to similar blocks n, is $e_1, e_2, ..., e_n$. Screen e block can be solved by weighted average of similar blocks. The solution formula is as follows:

$$\hat{e} = \sum_k w(d) e(d) \tag{2}$$

Expression: the d similar block and weight are described by $e(d)$ and $w(d)$. And the weights meet the following conditions:

$$\sum_d w(d) = 1, 0 \le w(d) \le 1 \tag{3}$$

$w(d)$ can be described by equation (4):

$$w(d) = \frac{1}{Z} \exp\left(-\frac{l(d)}{h^2}\right) \tag{4}$$

In the formula: the pixel intensity of the picture block is represented by h^2. The Euclidean distance and normalization constant of picture blocks e and e_k are $l(d)$ and Z respectively.

$$l(d) = \|e - e_k\|_G^2 \tag{5}$$

In the formula: the controllable coefficient of the picture block is represented by k. The grayscale of the high-frequency picture block is denoted by G. On this basis, Z is solved by formula (6):

$$Z = \sum_k \exp\left(-\frac{l(d)}{h^2}\right) \tag{6}$$

After the self-similarity super-resolution reconstruction of the AR live broadcast screen is completed by the above method, it will ensure the integrity of the detailed texture features and implement the self-similarity reconstruction.

In order to obtain the most similar matching block R in the vicinity of the retrieval window f. After the matching search is completed through e, the high-frequency picture block R_{mn} corresponding to R is obtained, and belongs to HF. The similarity indicator is described by the mean of the sum of absolute values of the difference between two picture blocks. Therefore, the total absolute value of the difference between the two picture blocks is the matching block with the highest degree of similarity [7], then:

$$R = \arg\min |e - f| \tag{7}$$

The calculation formula to obtain the high-frequency picture block of R_{mn} is:

$$R_{mn} = R - R^* C_0 \tag{8}$$

Where the guided filter and convolution are described by C_0 and $*$, respectively. To obtain the sum of $\hat{e}$ and R_{mn}, it is done by formulas (2) and (8). Its calculation formula is:

$$e = \hat{e} + R_{mn} \tag{9}$$

B. Improved Singular Value Decomposition Detail Texture Enhancement

The high-frequency picture block R_{mn} obtained above is processed as follows with the improved singular value decomposition:

1) Watermark embedding: Let $W_{pq}(p = q)$ denote the matrix with embedded watermark, and p denote the dimension of the matrix.

a) After the high-frequency picture block R_{mn} is decomposed by the first-order wavelet transform, a new matrix B can be obtained, and it is composed of low-frequency coefficients [8].

b) Use 16×16 blocks to perform processing on B, then $B_{ij} = f_{ij}(x, y)$ represents each block after processing, and $i = 1, 2, ..., m$, $j = 1, 2, ..., n$, $x \ge 1$, $16 \ge y$. The decomposition formula of each block matrix after processing is:

$$B_{ij} = U_{ij} S_{ij} V_{ij}^T \tag{10}$$

Where: U and V are orthogonal matrices.

c) Suppose the largest singular value of S_{ij} is denoted by φ_{ij}. According to the properties of singular value decomposition, φ_{ij} is adjusted, and the adjustment formula is:

$$if\left(W_{ij}=1\right)\begin{cases} if\left(\bmod\left(\varphi_{ij},a\right)<a/4\right)\varphi_{ij}= \\ \varphi_{ij}-\left(\bmod\left(\varphi_{ij},a\right)-a/4\right); \\ else\varphi_{ij}=\varphi_{ij}-\left(\bmod\left(\varphi_{ij},a\right)+3^{*}a/4\right) \end{cases}$$
$$else\begin{cases} if\left(\bmod\left(\varphi_{ij},a\right)>3^{*}a/4\right)\varphi_{ij}= \\ \varphi_{ij}-\left(\bmod\left(\varphi_{ij},a\right)+5^{*}a/4\right); \\ else\varphi_{ij}=\varphi_{ij}-\left(\bmod\left(\varphi_{ij},a\right)+a/4\right) \end{cases} \tag{11}$$

In the formula: $\bmod(x,y)$ is used to represent $x \bmod y$. The watermark signal values corresponding to the embedded strength factor and S_{ij} are a and W_{ij}, respectively. The adjusted S_{ij} is represented by $\widehat{S}_{ij}$, and singular value decomposition is performed on all $\widehat{S}_{ij}$, so that each block contains watermark information [9].

d) According to the ordinal position, all the blocks containing the watermark are recombined to form a new matrix R' of low frequency coefficients. Reconstruction method is used to process the changed wavelet coefficients [10] to obtain the super-resolution grayscale image I' with the initial watermark embedded.

2) Watermark extraction

a) I' is processed by first-order wavelet decomposition to obtain a matrix M, which consists of decomposed low-frequency coefficients.

b) Use 16×16 blocks to process M, then $M_{ij}=q_{ij}(x,y)$ represents each block after processing, and $i=1,2,...,m$, $j=1,2,...,n$, $x\geq 1$, $16\geq y$.

$$M_{ij}=U'_{ij}S'_{ij}V'^{T}_{ij} \tag{12}$$

c) If the largest singular value of S'_{ij} is represented by σ'_{ij}, and adjust it, its formula is:

$$iq\left(\bmod\left(\varphi'_{ij},s\right)>s/2\right)\left\{W_{ij}=1\right\}else\left\{W_{ij}=0\right\} \tag{13}$$

The above steps are taken, and each block is processed to obtain the restored watermark binary picture B_i.

Referring to the overlapping characteristics of the blocks, it is impossible to directly add the B_i obtained by reconstruction to M_{ij}, and a control strategy should be adopted for the overlap to avoid the interference caused by it [11].Therefore, the window processing is performed on the picture block, which is completed by the Gaussian function. And it is centrally symmetric, and $\bar{R}_{mns}$ is obtained after processing, and its formula is:

$$\bar{R}_{mns}=\bar{R}_{mn}G_{\varphi}\left(\left\|x-x_c\right\|\right) \tag{14}$$

Where $G_{\varphi}\left(\left\|x-x_c\right\|\right)=\frac{1}{2\pi\varphi^2}u^{-\frac{\left\|x-x_c\right\|^2}{2\varphi^2}}$ is the Gaussian window function.

The acquired $\bar{R}_{mns}$ and M_{ij} are combined to form a super-resolution picture M^{i}_{mn}, which contains all the detailed texture features of the high-resolution picture [12], which is:

$$M^{i}_{mn}=M_{ij}+\bar{R}_{mns} \tag{15}$$

C. Algorithm Flow

The implementation of adaptive enhancement algorithm based on approximate matching for AR live-broadcast image details texture is divided into two parts. One is to get the high frequency information of the best matching block based on approximate matching. Secondly, the improved singular value decomposition is used to reconstruct the matrix to obtain the effective high frequency texture information. The detailed texture is synthesized to obtain the enhanced ultra-detailed texture [13-15]. Its enhanced overall flow is shown in Fig. 1.

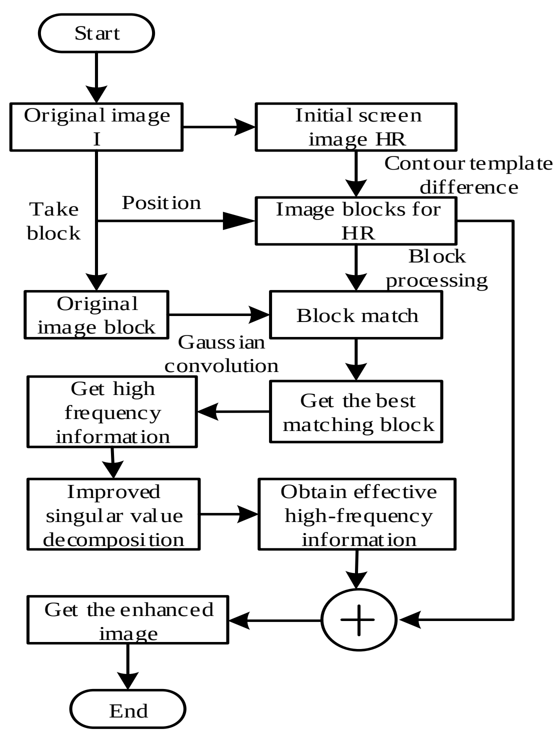

Fig. 1. Algorithmic flow.

III. Test Analysis

In order to test this algorithm for AR live picture details texture enhancement effect. Five live images are randomly selected from the image database of an AR live company in order to verify the effectiveness of the algorithm. The test environment operating system was Windows 10 x 64 8 GB of memory with CPU 3.4 GHz. The application development environment is Matlab 2018a.

In order to test the influence of scaling factor values on approximate matching, the sizes of matching blocks and retrieval windows are set to 2 × 2 and 4 × 4 respectively, and the gray space values (0 ~ 255) are used as features to obtain the normalized mean of the total absolute values of the difference between the two blocks. The smaller the value, the higher the similarity of the matching block. Test the results of the different zoom factor values, as shown in Fig. 2.

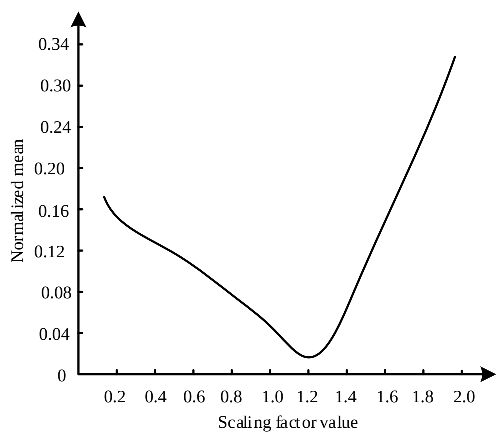

Fig. 2. Normalized mean change results.

According to the test results in Fig. 2, it can be seen that with the increase of the scaling factor value, the normalized mean value first decreases and then increases. When the scaling factor is 1.2, the normalized mean is the smallest. So in this article, determine that the zoom factor has a value of 1.2. In order to more intuitively reflect the impact of the algorithm with the scaling factor changes in image quality. We plotted the changes in the FSIM (feature similarity) metrics to show the picture quality, as shown in Fig. 3.

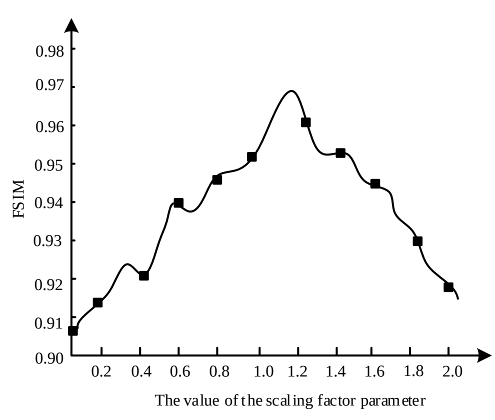

Fig. 3. Influence of Zoom Factor Parameter Value on Screen.

As shown in Fig. 3, the FSIM value is highest when the scale factor parameter is set to a value of 1.2. When the scaling factor parameter value reaches 0.97, the picture quality is the best. Therefore, the best scaling factor for this article is 1.2.

The quality of reconstructed image has some influence on the effect of texture enhancement. In order to measure the quality of the reconstructed images, peak signal-to-noise ratio (PSNR), structural similarity (MSSIM) and feature similarity (FSIM) are used as measurement criteria. The higher the value is, the better the quality is. The calculation formulas are as follows:

$$PSNR = 10\lg\left(\frac{255^2}{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X_1(i,j) - X_2(i,j)^2\right)}\right) \tag{16}$$

$$MSSIM = \frac{1}{W}\sum_{i=1}^{W}(b_i)^{\alpha}(c_i)^{\rho}(h_i)^{\gamma} \tag{17}$$

$$FSIM = \frac{\sum S_L(X_1, X_2)G(X_1, X_2)}{\sum G(X_1, X_2)} \tag{18}$$

Type: original picture and reconstruction picture for X_1 and X_2. The brightness, contrast and structure comparison functions of L block are b_i, c_i and h_i. The corresponding adjustment weights are α, ρ and γ. The number of blocks is W. The phase congruence is described by $S_L(X_1, X_2)$. The similarity measure of gradient features is described by $G(X_1, X_2)$.

In the process of decomposition, the value of the threshold has a great influence on the effect of image details texture enhancement. We choose PSNR and MSSIM as the metrics to get the effect of enhancing the image detail texture under different threshold values. The results are shown in Fig. 4 and Fig. 5.

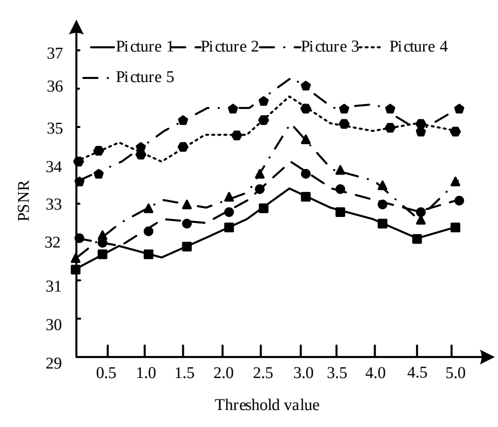

Fig. 4. Picture enhancement PSNR results with different thresholds

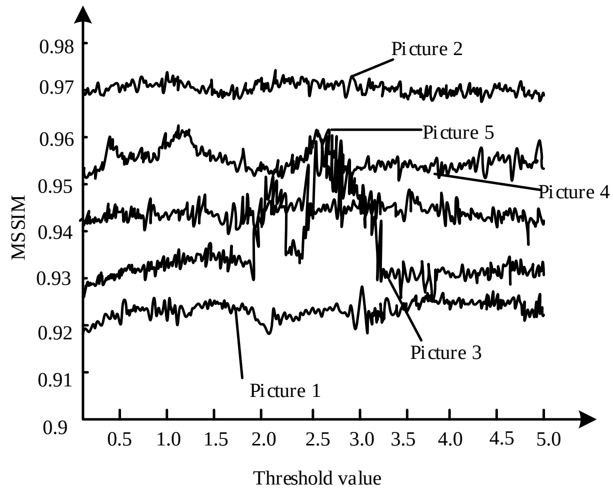

Fig. 5. Picture enhancement MSSIM results with different thresholds

According to Figs. 4 and 5, the PSNR and MSSIM results of the five live images are changed under different thresholds. When the threshold value is 3.0, the PSNR and MSSIM values of all live images after enhancement are the highest, indicating that the effect is best when the threshold value is 3.0. So the threshold value is 3.0 in the following experiment.

The enhanced algorithm based on the weight of the guide coefficient (Document [4] algorithm), the enhanced algorithm based on the multi-scale of NSST (Document [5] algorithm) and the enhanced algorithm based on the deep learning algorithm (Document [6] algorithm) are selected as the comparison algorithm. Through the above formula to obtain the effect of three kinds of algorithm picture reconstruction, as shown in Table I.

According to the test results in Table I, the results of PSNR, MSSIM and FSIM show that the three values of this algorithm are the best, significantly better than the three contrast algorithms. Therefore, it is shown that the reconstruction features of the proposed algorithm are closer to those of the original image and the reconstruction effect is better.

In order to test the subjective effect of the enhanced method, the detailed texture visual effect image of the enhanced image is obtained by three methods. Due to space constraints, only Screen 3 was selected for the test, as shown in Fig. 6.

According to the test results in Fig. 6, it can be seen that after the enhancement, the clarity and edge sharpness of the texture details are better than those of the three contrast algorithms. Experimental results show that the proposed algorithm is better than three contrast methods.

TABLE I. RECONSTRUCTION RESULTS OF THREE ALGORITHMS

Test content	Image number	Algorithm of this paper	Bootstrap weighted enhancement algorithm	Multiscale enhancement algorithm based on NSST	Enhancement algorithm based on deep learning
PSNR/dB	1	28	23	24	23
	2	29	25	26	24
	3	34	30	31	29
	4	32	29	27	28
	5	26	22	23	23
MSSIM	1	0.94	0.89	0.91	0.91
	2	0.91	0.86	0.85	0.87
	3	0.95	0.91	0.90	0.90
	4	0.93	0.87	0.88	0.86
	5	0.93	0.90	0.87	0.89
FSIM	1	0.93	0.89	0.87	0.88
	2	0.92	0.88	0.89	0.88
	3	0.98	0.93	0.92	0.92
	4	0.97	0.91	0.93	0.92
	5	0.94	0.89	0.90	0.91

(a) The enhancement effect of this algorithm

(b) Enhancement algorithm based on guide coefficient weighting

(c) NSST based multi-scale enhancement algorithm

(d) Enhancement algorithm based on deep Learning

Fig. 6. The visual comparison effect of the three methods.

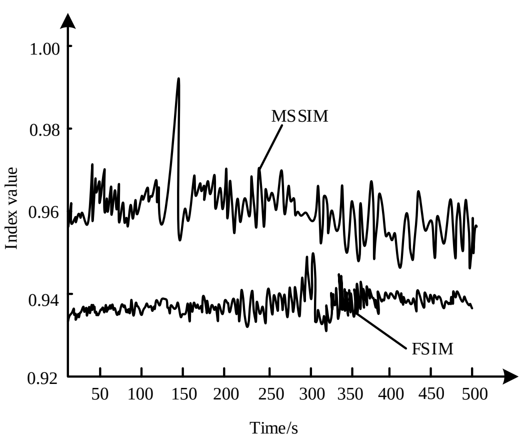

Fig. 7. Changes in MSSIM and FSIM indicators.

The algorithm designed in this paper is to improve the quality of AR live images, but the live images change with time. Therefore, it is necessary to verify the quality of the algorithm. With time as a variable, PSNR, MSSIM, and FSIM as metrics, the test results are shown in Fig. 7 and Fig. 8.

According to the changes of the three indexes in Fig. 7 and Fig. 8, the indexes of MSSIM and FSIM are more than 0.93 and PSNR is more than 26dB, and the indexes are higher. Therefore, the algorithm presented in this paper has good continuity effect and meets the requirements of AR live broadcast.

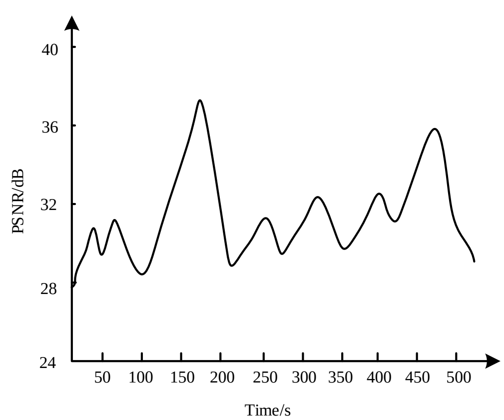

Fig. 8. Changes in PSNR indicators.

IV. Conclusion

AR live broadcast is the product of the rapid development of the Internet. It has been used in many fields, such as education, e-commerce and so on. This paper uses approximate matching algorithm to enhance the image texture details in order to improve the visual effect of AR live images. Self-similarity high frequency matching block is used to obtain and improve the singular value decomposition algorithm to reconstruct the image details and enhance the image texture. Experimental results show that the algorithm has good enhancement effect, and the sharpness and clarity of texture details are improved.

Acknowledgment

The study was supported by Scientific research project of the first-class Municipal Science and Technology Commission: Youth Fund of Beijing Academy of agricultural and Forestry Sciences, "recognition and production state determination of greenhouse vegetables based on image recognition" (qnjj202213).

References

[1] Cao S, Huang L P, Hou B B, et al. Self-Adaptive Non-Local Means Image Denoising Algorithm Based on Fuzzy Edge Complement[J]. Laser & Optoelectronics Progress,2018,55(01):213-218.

[2] Huang Y, Niu B N, Guan H, et al. Image texture based adaptive watermarking algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics,2019,45(12):2403-2414.

[3] Wu Y F, Zhu S C. Image Inpainting Algorithm Based on Matching Adjustment Rule and Gradient Constraint Model[J]. Packaging Engineering,2018,39(13):239-244.

[4] He L F, Zhou G B, Yao B, et al. A haze removal algorithm based on guided coefficient weighted and adaptive image enhancement method[J]. Microelectronics & Computer,2020,37(09):73-77+82.

[5] Wang X H, Zhai Y X, Ma X C. Retinex Low Light Image Enhancement Algorithm Based on Multi-Scale Adaptive NSST[J]. Packaging Engineering,2020,41(03):211-217.

[6] Xu S P, Lin Z Y, Zhang G Z, et al. A Low-Light Image Enhancement Algorithm Using the Hybrid Strategy of Deep Learning and I mage Fusion[J]. Acta Electronica Sinica,2021,49(1):72-76.

[7] Long X, He G T. Image enhancement method based on multi-layer fusion and detail recovery[J]. Application Research of Computers,2020,37(02):584-587.

[8] Lv B K, Wu C M, Tian X P. Fast Image Matching Algorithm Based on Best-Buddies Similarity[J]. Laser & Optoelectronics Progress,2020,57(10):189-195.

[9] Wang X C, Wang L L, Mo B, et al. Image Restoration Algorithm for Rotary Motion Blur Based on Adaptive Gradient Priori Regularization[J]. Acta Armamentarii,2019,40(03):522-529.

[10] Liu C A, Ai Z, Zhao L J. Self-adaptive image feature matching algorithm based on grid motion statistics[J]. Journal of Huazhong University of Science and Technology(Nature Science Edition),2020,48(01):37-40+54.

[11] Zhang C F, Shao L P, Shi J. Improved Criminisi Algorithm Based on Adaptive Gradient Classification and Matching[J]. Journal of Chinese Computer Systems,2018,39(02):379-385.

[12] Qin Z K, Zhang H Y. Adaptive texture image retrieval based on transform domain[J]. Chinese Journal of Liquid Crystals and Displays,2020,35(03):281-286.

[13] Wang H, Zhang S W, Xu K. A Real-Time Adaptive RANSAC Algorithm Applied in Image Matching[J]. Electronics Optics & Control,2020,27(02):90-93+97.

[14] Yin C, Guo X J, Tian X, et al. Contrast and detail enhancement based on structure and texture[J]. Application Research of Computers, 2019, 35(12):318-321.

[15] Yang H F, Li H, Chen K Y, et al. Image feature points extraction and matching method based on improved ORB algorithm[J]. Journal of Graphics,2020,41(04):548-555.