



Evaluating the robustness of image matting algorithm

ISSN 2468-2322

Received on 21st March 2020

Revised on 5th June 2020

Accepted on 18th June 2020

doi: 10.1049/trit.2020.0079

www.ietdl.org

 Genji Yuan^{1,2}, Jinjiang Li^{1,2} ✉, Hui Fan^{1,2}
¹School of Computer Science and Technology, Shandong Technology and Business University, Yantai 264005, People's Republic of China

²Co-innovation Center of Shandong Colleges and Universities: Future Intelligent Computing, Shandong Technology and Business University, Yantai 264005, People's Republic of China

✉ E-mail: lijjiang@gmail.com

Abstract: In this study, the authors propose a method to calculate the consistency of alpha masking to assess the robustness of the matting algorithm. This study evaluates consistent alpha masks based on the Gaussian–Hermite moment in combination with gradient amplitude and gradient direction. The gradient direction describes the appearance and shape of local objects in the image, and the gradient amplitude accurately reflects the contrast and texture changes of small details in the image. They selected Gaussian blur, pretzel noise, and combined noise to destroy the image, and then evaluated the consistency of the original alpha mask and noise alpha mask. To determine the robustness of the matting algorithm, they assessed the degree of consistency of the alpha mask using three different evaluation levels. The experimental results show that noise has a greater impact on the performance of the matting algorithm, which shows a decreasing trend as the noise level in the image deepens. In noisy images, the traditional matting algorithm exhibits better robustness compared to the recently proposed trap matting algorithm. Different matting algorithms present different adaptations to different noises.

1 Introduction

Natural image matting is the process of accurately estimating the unknown region between a user-defined foreground object and a background image. The opacity of each pixel is precisely defined in the process of estimating the soft partition. And the alpha mask estimated by the matting algorithm is one of the core elements in image and video editing work. In image and video compositing, alpha masks are essential. The foreground object is accurately extracted through the alpha mask, and then it is composited with the new background to render the new scene. Formally, assuming that the observed image I consists of three parts: the foreground object F , the background image B , and the alpha mask α , the mixed colours in the unknown region between the foreground and background can be represented using the following model:

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i, \quad (1)$$

where $\alpha_i \in [0, 1]$ represents the opacity of the foreground object at a pixel i . Since the pixel values of the foreground and background are not known, the matting problem is a serious under-constrained problem, containing seven unknowns and three known quantities. Typically, the user provides prior information for each pixel in the form of trimap. The trimap contains the foreground region ($\alpha = 1$), the background region ($\alpha = 0$), and the unknown region. The matting algorithm aims to estimate the unknown alpha value using the known pixel colour. The key task of natural image matting is to determine the alpha value of the mixed pixels. For the ideal matting algorithm, the alpha value of the mixed pixel can be accurately estimated under different backgrounds and different interference factors. Currently, matting algorithms are estimating alpha masks on a clean and clear image. No algorithm attempts to estimate alpha masks on images containing noise. These matting algorithms use multiple quality assessment methods to evaluate their estimated alpha masks. However, these evaluation methods only evaluate the performance of the matting algorithm and not the robustness of the matting algorithm. In this paper, we propose

a new evaluation method that aims to assess the robustness of the matting algorithm. In this paper, the alpha masks estimated by different matting algorithms in common noisy images are evaluated using Gaussian–Hermite moments [1, 2]. Gaussian–Hermite moments can extract image features well and thus estimate the consistency of alpha masks.

Different matting algorithms have different adaptations for inferior images or different backgrounds. The same matting algorithm may yield different results when estimating the alpha mask of the same foreground object in a different context. Because of the different background information, the MATTING algorithm gets a large difference in alpha masking even when estimating the same foreground objects. And in practice, the images we obtain may not be clear. The robustness of the matting algorithm can be demonstrated by the consistency of the alpha mask. The quality evaluation of images [3–5] is mainly divided into subjective and objective evaluations [6]. Subjective evaluations are mainly observed and evaluated directly by the human eye, consistent with the perceptual characteristics of the person, but susceptible to factors such as subjective emotions. Objective evaluations can be divided into fully referenced evaluations, partially referenced evaluations and non-referenced evaluations [7–9]. Full reference evaluation is the evaluation of image quality when the reference image is fully available, and partial reference evaluation is the evaluation of an image using part of the information from the reference image. No-reference evaluations [10] are performed directly on the image without any information from the reference image. Among them, the results of the full reference evaluation are relatively reliable and stable and are called the most reliable evaluation algorithm.

Mean square error (MSE) and peak-signal-to-noise ratio (PSNR) are commonly used methods for image quality evaluation. However, these two methods do not consider the correlation between image pixels when evaluating image quality, and their evaluation results are less consistent with the subjective evaluation. Structural similarity (SSIM) [11], based on the theory that the human eye can extract structural information from images, uses mean, variance and covariance to measure the local brightness,

contrast and structural similarity of images, and thus the overall similarity of images, and the evaluation results are more consistent with the human subjective perception. When evaluating alpha masks, the key is the variation of pixels within the unknown region. The foreground and background regions defined in the alpha mask, even with the addition of noise, do not strongly influence the final fusion results. The unknown region tends to be the edge of the foreground object with a more drastic gradient change. The gradient is an important component of the boundary. The gradient direction can represent the structural information of the boundary and can accurately describe the local shape information of the boundary. In this paper, the consistency of the alpha mask and thus, the robustness of the matting algorithm is evaluated using gradient amplitude [12] and gradient direction.

In this paper, the robustness of different matting algorithms is evaluated under Gaussian blur [13], pretzel noise and combined noise. In the evaluation process of this paper, not only the matting algorithms are compared under different noise conditions, but also the different matting algorithms are compared. In order to quantitatively assess the effect of noise on the performance and results of the matting algorithm, we also used MAD and RMSE [14] to evaluate the matting results. The main objective of this paper is to assess the effect of noise on the matting algorithm and to observe the quality of the alpha masks estimated by different matting algorithms on the noise image to assess the robustness of the matting algorithm. In the matting algorithm [15], it is difficult to accurately estimate the alpha value of foreground objects, especially for estimating fine hairs and translucent objects. Adding noise to the image increases the difficulty of the matting algorithm to estimate alpha masks. Based on the evaluation method in this paper, we estimate the robustness of the matting algorithm by the estimated alpha value of the mixed pixels. The better the robustness of the matting algorithm, the more consistent alpha masking it can estimate on the noise image.

A Gaussian blur filter can blur the high-frequency details of an image so that specific detail information is reduced. Pretzel noise (pepper and salt) can cause randomly distributed different noises on the image, each independent and uncorrelated. When both noises are used at the same time, Gaussian blur can interfere with some specific details in the image, and pulsed noise can randomly diffuse pixels in the image, both destroying pixels in specific areas of the image. We first used a Gaussian blur filter and pulsed noise to destroy the image. The alpha mask is then re-estimated on the noise-containing image using the matting algorithm. The noise we use contains four different levels. During the experiment, the Gaussian-Hermite moment was used to assess the similarity between the estimated matting results at the four noise levels and the original matting results. We evaluate the robustness of the matting algorithm at different noise levels. We set three evaluation levels (i.e. 0.3, 0.6 and 0.8) to more rigorously assess the robustness of the matting algorithm. In all noise levels, these three thresholds serve as the basic assessment conditions. This paper evaluates the robustness of the matting algorithm using Gaussian-Hermite moments combined with gradient amplitude and gradient direction [16], and this paper verifies the robustness of different matting algorithms on fuzzy data sets.

The remaining of this paper is organised as follows. In Section 2, we briefly review the classical matting algorithm and explain the basic terminology. In Section 3, we describe the algorithm of this paper in detail. In Section 4, we discuss the experimental results of this paper in detail. In Section 5, we provide a comprehensive summary of the paper and look forward to future work.

2 Related work

Natural image matting algorithms [17, 18] can be mainly divided into three categories: sampling-based matting algorithms [19], alpha propagation-based matting algorithms and deep learning-based matting algorithms.

The sample-based approach inferred the alpha value of pixels in unknown regions of the image by taking known pixel colours. The

key task of the sampling-based matting algorithm is to sample the pixels and construct a foreground colour and background colour model based on the captured pixels. The sampling-based matting algorithm uses the idea of image statistics to solve the matting problem. When the accuracy of trimap is high enough, the pixel colours of the unknown region can be estimated from the constructed model, making the colour distribution of the unknown region strongly correlated with the foreground and background. Bayesian matting [20] estimated the alpha value of the unknown region based on local colour statistics. Bobust matting [21] tended to collect spatially close pixel pairs from a known background region of the foreground. Shared matting [22] argued that the real sample might be farther away from the unknown pixels, and they collected samples from the boundary of trimap. Global matting [23] sampled all pixels at the boundary of trimap and collected a large number of samples. While this greatly reduces the probability of missing the true sample, it also increases the computational cost. Comprehensive sampling [24] does not sample directly from the trimap boundary, but instead expresses the sampled distance as a variable based on the distance of the unknown pixel from the trimap boundary to determine the sampling distance. Karacan *et al.* [17] proposed a sparse sampling approach, where they used KL-divergence to assess similarity at the super-pixel level to determine the sampling distance. The method aims to find two independent samples for each pixel in the unknown region.

The alpha-based propagation method propagates the alpha value of a known foreground or background pixel into an unknown region based on a similarity function. The similarity function is generally defined as spatial proximity and colour similarity. Poisson matting [25] estimated alpha masking by solving the Poisson equation. Levin *et al.* [26] proposed a closed-form approach to solve the matting problem. They used local colour statistics to estimate the affinity between two pixels. Later, He *et al.* [27] proposed to solve the matting problem using a large kernel. Based on the non-local principle, Lee and Wu [28] proposed a non-local matting method, while KNN matting [29] argues that each pixel and their non-local pixels have similar alpha values. The information-flow matting [30] achieves high-quality alpha masking by combining local and non-local correlations.

The deep learning-based matting method learns the mapping of alpha masks directly from the input image. Cho *et al.* [31] proposed an end-to-end CNN network based on the closed matting algorithm and the KNN matting algorithm. Xu *et al.* [32] used the image and the corresponding trimap as input to learn alpha masks using an encoder-decoder network. Finally, using a small convolutional network to optimise the alpha mask. Lutz *et al.* [33] applied generative adversarial networks (GANs) to image matting. They used the atrous spatial pyramid pooling module to improve the encoder-decoder network and feature sampling at multiple scales.

Wee *et al.* [34] proposed an image quality assessment method based on discrete orthogonal moments. The method divides the image into image blocks, then calculates the discrete orthogonal matrices of the individual image blocks, and calculates the quality fraction of the image blocks using the correlation coefficients of the moment transformation matrix corresponding to the reference image and the test image. The information fidelity algorithm (IFC) [35] proposes an algorithmic framework for combining image quality with visual perception based on information theory. The visual information fidelity (VIF) method [36] then argues that the quality of the test image can be quantified by the amount of information lost by the test image relative to the reference image. The image quality evaluation method based on the underlying features (FSIM) [37] uses phase consistency to assess image quality, but there are limitations in detecting some of the more acute image features such as step edges. The gradient modulus similarity deviation (GMSD) algorithm [12] was designed from the degree of retention of the structural features and the pooling strategy two-sided algorithm. Use gradients as effective primary visual features to describe the image contrast. The boundary strength similarity (ESSIM) algorithm [38] assumes that perceived boundaries in images are the basis for semantic cognition. The

boundary features extracted by this algorithm are usually distributed in non-local regions of the image. Li *et al.* [11] combined weighted gradient information with SSIM to effectively improve the accuracy of quality assessment. Lee *et al.* [39] improved the structural comparison function of SSIM using the standard deviation correlation function of mean segmentation and introduced the sharpness comparison function considering the effect of image sharpness.

Noise [40, 41] often manifests itself on images as pixel dots or pixel blocks that cause stronger visual effects, disrupting the observability of the image. Noise is generally generated during the acquisition of an image or during the transmission of an image signal. The distribution and magnitude of the noise across images are irregular and random in nature. There is a general correlation between noise and image. Different noises can be superimposed between them. The pulsed noise will randomly change some pixel values, producing a black and white bright and dark noise. Pretzel noise is often caused by image cutting. From a mathematical point of view, the Gaussian blur of an image is the image doing convolution with the Gaussian distribution. Since the Fourier variation of the Gaussian function is another Gaussian function, the Gaussian blur is a low-pass filter for the image.

3 Proposed method

We use different matting algorithms to estimate alpha masks from the original and noisy images, respectively. We evaluated the similarity between the original alpha mask and the noise alpha mask by combining gradient amplitude similarity and gradient direction similarity [42] with the Gaussian–Hermite moment. The reference image and the test image were first divided into non-overlapping chunks [43]. The gradient magnitude and direction of the gradient were then calculated for the reference and test images, respectively. Second, the gradient amplitude similarity and gradient direction similarity between the test image and the reference image were calculated. After that, the standard deviation of the statistical gradient amplitude similarity and the standard deviation of the gradient direction similarity for all image blocks. Finally, the moment energy differences of the image blocks were calculated using gradient amplitude similarity and gradient direction similarity [44] to obtain the consistency assessment of the alpha mask.

We use different levels of noise to destroy the original image and compare the noise alpha mask to the original alpha mask. For consistency assessment, we treat the original alpha mask as ground truth to calculate the consistency factor for the noise alpha mask. We evaluate the robustness of the matting algorithm across the entire data set. We use ground truth alpha masks to blend different background images to increase the number of images in the dataset.

3.1 Gaussian–Hermite moment

The theoretical definition interval of the Gaussian–Hermite matrix is $(-\infty, +\infty)$. Assuming that the size of the input image I is $K \times K$, in order to calculate the moment transformation matrix of the input image, we convert the kernel function $H_p(x/\sigma)$ of the Gaussian–Hermite moment into the discrete kernel function $\tilde{H}_p(i, K; \sigma)$. The conversion formula is as follows:

$$\tilde{H}_p(i, K; \sigma) = (2^p p! \sqrt{\pi} \sigma)^{-1/2} \exp(-X^2/2\sigma^2) H_p(X/\sigma). \quad (2)$$

Equation (2) is the discrete form of the Gaussian–Hermite moment kernel function, and the moment transformation of the image can be expressed as

$$\eta = \frac{4}{(K-1)^2} \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} I(i, j) \tilde{H}_p(i, K; \sigma) \tilde{H}_q(j, K; \sigma), \quad (3)$$

where $p+q$ is the order number and $p \in [0, K-1]$, $q \in$

$[0, K-1]$. Because of the orthogonality of the orthogonal moment, the image can be reconstructed by using the inverse transformation of (3)

$$I(i, j) = \sum_{p=0}^n \sum_{q=0}^n \eta_{p,q} \tilde{H}_p(i, K; \sigma) \tilde{H}_q(j, K; \sigma). \quad (4)$$

The orthogonal moments have orthogonal and invariant properties and can be used to extract edge features of the image.

The main differences in alpha masking estimated by different matting algorithms exist at the edge of the foreground object, and the gradient is an important component of the boundary. The gradient amplitude can reflect changes in the detail of the alpha mask. We use gradient amplitude similarity to measure the degree of distortion in noise alpha.

3.2 Image degradation

Gaussian blur is an image blur filter that uses a normal distribution to calculate the transformation of each pixel in an image [45]. Gaussian blur smoothes out the high-frequency information of an image and reduces its specific details. After processing the image using Gaussian blur, the gradient value of the image boundary decreases dramatically [46]. The Gaussian function in two-dimensional space is defined as

$$G(u, v) = \frac{1}{2\pi\sigma^2} e^{-(u^2+v^2/2\sigma^2)}, \quad (5)$$

where r is the blur radius and σ is the standard deviation of the normal distribution. u, v represent the distance between the horizontal axis and the vertical axis and the origin, respectively. The Gaussian filter, when applied to image processing, produces a concentric circle with a normal distribution from the centre. A convolutional matrix of pixels whose distribution is not zero is convolved with the original image, and the values of each pixel are a weighted average of the values of neighbouring pixels. The values of the original pixels have the largest Gaussian distribution values and therefore have the largest weights. Adjacent pixels are given less and less weight as they get further away from the original pixel.

Pretzel noise [45] is one of the common noises in image processing. Pretzel noise is the bright black and white dark spots on the image caused by image sensors, transmission channels, decoding processing, etc. Pretzel noise includes salt noise (greyscale –255) and pepper noise (greyscale –0), the former being high greyscale noise and the latter being low greyscale noise. It is a randomly distributed noise that may appear as black pixels in bright areas, white pixels in dark areas, or both. Fig. 1 shows the noise image used in this paper. We used four different levels of noise to destroy the original image (Table 1).



Fig. 1 Noise image. The first row is a Gaussian blurred image, the second row is pretzel noise image, and the third row is a combination of image blur and pretzel noise image. We divide the noise into four different levels

Table 1 RMSE and MAD evaluation results (Gaussian blur)

	original		1 blur		2 blur		3 blur		4 blur	
	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD
deep	2.4834	1.5838	2.6920	2.0586	2.9678	2.3673	3.1134	2.7457	3.1946	2.9474
KNN	2.6205	1.8256	2.6992	2.0361	2.8615	2.3691	3.0183	2.7269	3.1332	3.0220
LB	2.5683	1.3154	2.7173	1.4790	2.8770	1.7392	3.0083	1.9792	3.1118	2.1719
three	2.5044	1.3253	2.7084	1.5753	2.8749	1.8937	3.0027	2.1443	3.0563	2.2789
WCT	2.7302	1.7838	2.9018	2.0043	3.0577	2.3741	3.1813	2.6663	3.2732	2.8840

Bold indicates the best performance.

Algorithm 1: Robustness assessment algorithm based on the Gaussian-Hermite moment.

Input:

Original image I , Gaussian blur filter, pretzel noise, a combination of Gaussian blur and pretzel noise;

Output:

Evaluation results Q ;

1. Estimate alpha masks on the original image I ;
2. Add noise (Gaussian blur, pretzel noise, and combined noise) to the original image Q by adding up to four levels of noise. Estimate alpha masking on a noise image N .
3. Using the Prewitt operator h_x in the horizontal direction and the Prewitt operator h_y in the vertical direction to calculate the gradient of any pixel point i in the image in the horizontal direction $G_x(i)$ and the vertical direction $G_y(i)$ and the magnitude of the gradient $G(i)$.

$$h_x = \begin{bmatrix} 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \end{bmatrix}, \quad h_y = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 \\ -1/3 & -1/3 & -1/3 \end{bmatrix},$$

$$G_x(i) = I(i) \otimes h_x,$$

$$G_y(i) = I(i) \otimes h_y,$$

$G(i) = \sqrt{G_x^2(i) + G_y^2(i)}$. The similarity of the gradient magnitude between the original alpha and the noise alpha can be expressed as

$$H(i) = \frac{2G_r(i)G_n(i) + C}{G_r^2(i) + G_n^2(i) + C}.$$

4. Calculate gradient direction similarity

$$T(i) = B(i)L(i),$$

$$B(i) = \frac{2m_r(i)m_n(i) + C_1}{m_r^2 + m_n^2 + C_1},$$

$$L(i) = \frac{2s_r(i)s_n(i) + C_2}{s_r^2 + s_n^2 + C_2}.$$

5. Divide the original alpha mask and noise alpha mask into 8×8 image blocks, and then calculate the low-order moment transformation matrix of the image blocks. Calculation of the moment energy difference of the image block from the energy value of the orthogonal matrix

$$D_k = (d_{ij})_{3 \times 3} = \begin{pmatrix} 2p_{ij}q_{ij} \\ p_{ij}^2 + q_{ij}^2 \end{pmatrix}_{3 \times 3}.$$

6. Calculate the correlation matrix by combining gradient amplitude similarity and gradient direction similarity for image blocks

$$I_k = \left(f(d_{ij}) \right)_{3 \times 3} = \begin{pmatrix} I_{00}^k & I_{01}^k & I_{02}^k \\ I_{10}^k & I_{11}^k & I_{12}^k \\ I_{20}^k & I_{21}^k & I_{22}^k \end{pmatrix},$$

$$f(d_{ij}) = d_{ij} \cdot H \cdot T.$$

7. Average the correlation matrix for all image blocks

$$I_k = \frac{1}{K} \begin{pmatrix} \sum_{k=1}^K I_{00}^k & \sum_{k=1}^K I_{01}^k & \sum_{k=1}^K I_{02}^k \\ \sum_{k=1}^K I_{10}^k & \sum_{k=1}^K I_{11}^k & \sum_{k=1}^K I_{12}^k \\ \sum_{k=1}^K I_{20}^k & \sum_{k=1}^K I_{21}^k & \sum_{k=1}^K I_{22}^k \end{pmatrix} = \begin{pmatrix} I_{00} & I_{01} & I_{02} \\ I_{10} & I_{11} & I_{12} \\ I_{20} & I_{21} & I_{22} \end{pmatrix}.$$

8. The results of the Alpha mask consistency assessment can be expressed as

$$Q = \frac{1}{9} \sum_{i=0}^2 \sum_{j=0}^2 \frac{I_{ij}}{(0.8n + 0.2m)/K}.$$

3.3 Robustness assessment

We first estimate the alpha mask of the original image based on different matting algorithms. We then applied the noise to the original image and estimated the alpha mask on the noise image.

The alpha mask of the original image is considered a ground truth image. We used the Gaussian-Hermite moment to evaluate the consistency of different alpha masks. We use the same matting algorithm to estimate alpha masks on the original and noisy images, respectively, whose unknown region alpha values are the object of our focus. We evaluate the consistency of the noise alpha mask with the original alpha mask at different noise levels. Based on the magnitude of the Gaussian-Hermite moment evaluation results, we can determine the robustness of different matting algorithms. The specific implementation process of the Gaussian-Hermite moment-based robustness evaluation algorithm in this paper is described in Algorithm 1.

The gradient is usually calculated using linear filter convolution. Classical filters include Roberts, Sobel, Prewitt, Scharr, and others. We use the horizontal Prewitt operator h_x and the vertical Prewitt operator h_y to calculate the gradient $G_x(i)$ in the horizontal direction and the gradient $G_y(i)$ in the vertical direction of any pixel point i in the image. Then we can get the gradient magnitude $G(i)$ of the pixel, where h_x, h_y are, respectively,

$$h_x = \begin{bmatrix} 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \end{bmatrix}, \quad h_y = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 \\ -1/3 & -1/3 & -1/3 \end{bmatrix}. \quad (6)$$

$G_x(i), G_y(i), G(i)$ can be expressed as

$$\begin{aligned} G_x(i) &= I(i) \otimes h_x, \\ G_y(i) &= I(i) \otimes h_y, \\ G(i) &= \sqrt{G_x^2(i) + G_y^2(i)}. \end{aligned} \quad (7)$$

Let the gradient amplitudes of the original alpha and noise alpha be G_r and G_n , respectively, then the gradient magnitude similarity can

be expressed as

$$H(i) = \frac{2G_r(i)G_n(i) + C}{G_r^2(i)G_n^2(i) + C}, \quad (8)$$

where C is a normal number. The gradient magnitude similarity is calculated based on the pixel method, and the gradient amplitude is calculated based on the local block. When G_r and G_n are the same, the value of $G(i)$ reaches a maximum of 1.

After adding noise to the original image, the alpha mask estimated by the matting algorithm is prone to distortion. Noise-induced image distortion can be easily observed. The gradient direction [46] can well describe the edge distortion of the alpha mask. Therefore, we use the gradient direction as another important measure of alpha masking.

We first calculate the horizontal gradient, vertical gradient, gradient amplitude and gradient direction of the original alpha mask and noise alpha mask. We divided the original alpha mask and the noise alpha mask into 8×8 image blocks. We calculate the gradient direction for each image block. Then, the gradient directions of all the image blocks in the image are connected to get the feature vectors of the gradient directions. We can get the gradient direction feature of the whole image. The gradient directions of the original alpha mask and noise alpha mask are expressed as D_r and D_n , respectively. The gradient directions are assumed to have a range of $0^\circ - 180^\circ$, we divided it into nine parts and voted on using a weighted gradient magnitude.

The similarity of the gradient direction is mainly determined by the mean and variance of the gradient direction, namely

$$\begin{aligned} T(i) &= B(i)L(i), \\ B(i) &= \frac{2m_r(i)m_n(i) + C_1}{m_r^2 + m_n^2 + C_1}, \\ L(i) &= \frac{2s_r(i)s_n(i) + C_2}{s_r^2 + s_n^2 + C_2}, \end{aligned} \quad (9)$$

where $m_r(i)$, $m_n(i)$ are the mean values of the gradient directions of the i th image block in the original alpha image and the noise alpha image, respectively, and $s_r(i)$, $s_n(i)$ are the variances of the gradient directions of the i th image block in the original alpha image and the noise alpha image, respectively. $T(i)$ represents the gradient direction similarity of the noise alpha. When the original alpha and the i th image block in the noise alpha are the same, $T(i)$ reaches the maximum value of 1.

Continuous orthogonal matrices have a good ability to extract image information. We first evaluate the consistency of the image blocks using the Gaussian-Hermite moment, then incorporate gradient amplitude similarity and gradient direction similarity, weight the blocks for different regions of the image, and finally take the average as the evaluation score of the image.

We divide the original alpha mask and noise alpha mask into image blocks of size 8×8 [47], and then calculate the low order moment transformation matrix of the image block, namely

$$A_k^r = \begin{pmatrix} a_{00}^r & a_{01}^r & a_{02}^r \\ a_{10}^r & a_{11}^r & a_{12}^r \\ a_{20}^r & a_{21}^r & a_{22}^r \end{pmatrix}, \quad A_k^n = \begin{pmatrix} a_{00}^n & a_{01}^n & a_{02}^n \\ a_{10}^n & a_{11}^n & a_{12}^n \\ a_{20}^n & a_{21}^n & a_{22}^n \end{pmatrix}, \quad (10)$$

where A_k^r , A_k^n represent the moment transformation matrices of the k th image block in the original alpha mask and noise alpha mask, respectively. Because the value of A_k^r , A_k^n may be non-positive, you cannot use A_k^r , A_k^n directly for related calculations. We make the following adjustments to A_k^r , A_k^n :

$$\begin{aligned} P_k &= (p_{ij})_{3 \times 3} = \text{abs}(A_k^r), \\ Q_k &= (q_{ij})_{3 \times 3} = \text{abs}(A_k^n), \end{aligned} \quad (11)$$

where P_k , Q_k are the energy values of the orthogonal matrix, and the moment energy difference of the image block can be obtained according to P_k , Q_k :

$$D_k = (d_{ij})_{3 \times 3} = \left(\frac{2p_{ij}q_{ij}}{p_{ij}^2 + q_{ij}^2} \right)_{3 \times 3}, \quad (12)$$

where $d_{ij} \in (0, 1]$, $i, j = 0, 1, 2$. Using the gradient magnitude similarity and gradient direction similarity of the image block, we can be obtained as follows:

$$f(d_{ij}) = d_{ij} \cdot H \cdot T, \quad (13)$$

where $h \in (0, 1]$, $t \in (0, 1]$, we can get the correlation matrix as

$$I_k = (f(d_{ij}))_{3 \times 3} = \begin{pmatrix} I_{00}^k & I_{01}^k & I_{02}^k \\ I_{10}^k & I_{11}^k & I_{12}^k \\ I_{20}^k & I_{21}^k & I_{22}^k \end{pmatrix}. \quad (14)$$

Average the correlation matrix of all image blocks

$$I_k = \frac{1}{K} \begin{pmatrix} \sum_{k=1}^K I_{00}^k & \sum_{k=1}^K I_{01}^k & \sum_{k=1}^K I_{02}^k \\ \sum_{k=1}^K I_{10}^k & \sum_{k=1}^K I_{11}^k & \sum_{k=1}^K I_{12}^k \\ \sum_{k=1}^K I_{20}^k & \sum_{k=1}^K I_{21}^k & \sum_{k=1}^K I_{22}^k \end{pmatrix} = \begin{pmatrix} I_{00} & I_{01} & I_{02} \\ I_{10} & I_{11} & I_{12} \\ I_{20} & I_{21} & I_{22} \end{pmatrix}, \quad (15)$$

where K is the total number of all image blocks in the image. Assuming that the number of image blocks containing an unknown area in the image is n , and the number of image blocks not including the unknown area is m , the evaluation result Q of the image can be expressed as:

$$Q = \frac{\frac{1}{9} \sum_{i=0}^2 \sum_{j=0}^2 I_{ij}}{(0.8n + 0.2m)/K}. \quad (16)$$

We evaluate the consistency of the noise alpha with the original alpha on the dataset based on the Gaussian-Hermite moment, which has an evaluation range of $[0, 1]$. A value of 0 indicates that there is no similarity between the two alpha masks, and a value of 1 indicates that the two alpha masks are definitely similar. Different thresholds were used to obtain different evaluation results when performing the evaluation on the dataset [48, 49]. We used three thresholds, using 0.3, 0.6 and 0.8, to evaluate different alpha masks. When there is a 30% similarity between the two alpha masks, the similarity index is 0.3. Similarly, for thresholds of 0.6 and 0.8, the similarity between alpha masks should be 60 and 80%. When the threshold is set to 0.3, alpha masks with thresholds greater than or equal to 0.3 are arranged in descending order for the dataset. Repeat the process at each noise level. Higher thresholds allow a rigorous assessment of the robustness of the algorithm.

We graded the alpha mask and evaluated the robustness of the matting algorithm under different conditions, depending on the defined threshold value. After calculating the similarity index of the alpha mask, it is indexed according to different thresholds to evaluate the performance of the matting algorithm. We index downward starting with the maximum similarity index. We evaluate the performance of the matting algorithm at different noise levels.

To better evaluate the performance of the matting algorithm, we also used MAD, RMSE to evaluate the experimental results of this paper. The smaller the values of MAD and RMSE, the better the predicted alpha mask quality.

4 Experiment

We experimented separately on three different noisy images. Experiment 1: Processing images using Gaussian blur. Experiment 2: Using pretzel noise to destroy images. Experiment 3: A combination of Gaussian blur and pretzel noise was applied to the original image. We evaluate the performance of the matting algorithm on these three different blurry images to obtain more robust results. We repeated these three experiments using the same thresholds. First, set the threshold value as $\theta = 0.3$, which is a low similarity. Second, let the threshold be $\theta = 0.6$, which is the intermediate similarity. Finally, the threshold value is set as $\theta = 0.8$, which is the most strict similarity. When $\theta = 0.3$, the performance of the matting algorithm is evaluated at a lower level of similarity. When $\theta = 0.6$, the performance of the matting algorithm is evaluated at an intermediate level similarity. When $\theta = 0.8$, the performance of the matting algorithm is evaluated at a strict level of similarity. We fuse different background images to create more fused images based on ground truth alpha masks. We evaluate the performance of the matting algorithm using publicly available benchmark datasets. We extended each foreground object to images with 200 different backgrounds, and we evaluated the robustness of the matting algorithm on a total of 5400 images. Fig. 2 shows the synthetic dataset of this paper. To achieve a more accurate quantitative assessment, we also used MAD and RMSE to evaluate our experimental results. We calculated the average MAD and RMSE for all alpha masks.

Fig. 3 shows the alpha mask estimated on the Gaussian blur noise image. In the lower part of Fig. 3, we show the consistent pixels of the original alpha mask with the noise alpha mask estimated in the unknown region. From this, it can be derived that the more severe the Gaussian blur, the fewer pixels of agreement between the estimated noise alpha masks and the original alpha masks. Particularly for pixels in the marginal hair portion, the pixels of the noisy alpha mask were rapidly reduced in agreement with the original alpha mask. In Fig. 4, we give the specific evaluation results for each matting algorithm. The number of consistent alpha masks for each matting algorithm at different evaluation levels is shown in Fig. 4. On the whole, the number of consistent alpha masks decreased significantly as the level of blur increased. Fig. 5 shows a comparison of the combinations of different matting algorithms at different assessment levels. From this, it can be concluded that the WCT matting algorithm exhibits the best performance with the best robustness. The LB algorithm has poor performance and does not get good results in Gaussian fuzzy datasets. Among them, the KNN matting algorithm is the most unstable, with Gaussian fuzzing having the greatest impact.



Fig. 2 Dataset synthesised in this paper. We use the ground truth alpha mask to blend foreground objects with different backgrounds to form new images

To better observe the effect of Gaussian blur on the performance of the matting algorithm, we present the quantitative evaluation results in Fig. 6. In the MAD quantitative evaluation, the LB matting algorithm has the best performance, followed by the Three matting algorithm. The Deep matting algorithm has the best performance in the RMSE quantitative evaluation, but Gaussian blurring has a greater impact. The performance of the deep matting algorithm decreases significantly as the level of ambiguity increases. From the quantitative evaluation of MAD and RMSE, it is known that Gaussian blurring has less impact on the Three matting algorithm. However, the three matting algorithm did not obtain the largest number of consistent alpha masks in the consistent alpha masking assessment presented in Fig. 5. From the evaluation of the Three matting algorithm presented in Fig. 4, it can be seen that the number of consistent alpha masks of the intermediate grade decreases rapidly with the enhancement of Gaussian blur. The number of consistent alpha masks for the three matting algorithm was able to remain essentially stable across the loose and strict evaluation ratings.

Fig. 7 shows the alpha mask estimated in the pretzel noise image. The first row in Fig. 7 shows the alpha mask estimated on the original image. Rows two through five show the alpha masks estimated in the four levels of noise images. The last four rows show pixels where the noise alpha mask is consistent with the original alpha mask in unknown regions. From the results presented in Fig. 7, the effect of pretzel noise on the matting algorithm is small. Most of the pixels at the edge of the image can be kept consistent. The amplification in Fig. 7 demonstrates the edge consistency of the alpha mask estimated by the matting algorithm at different levels of pretzel noise. Different matting algorithms have different adaptations to pretzel noise. From the results presented in Fig. 7, it can be concluded that pretzel noise has a greater effect on the Deep matting algorithm (Table 2). Most of the unknown pixels estimated by the original alpha mask and the noise alpha mask are not consistent. Whereas pretzel noise has less effect on the LB matting algorithm, the alpha mask estimated by the LB matting algorithm in the original image and the alpha mask estimated on the noise image are largely able to be consistent.

Fig. 8 shows the results of the specific evaluation of the matting algorithm for each test. From this, it can be concluded that low-level pretzel noise has less effect on the deep matting algorithm and LB matting algorithm. The alpha mask estimated by the matting algorithm can maintain good consistency even under the threshold conditions of strict evaluation. Whereas the KNN matting algorithm, the three matting algorithm and the WCT matting algorithm exhibit large differences in the different assessment thresholds. However, the amount of consistent alpha masking remained almost constant at different levels of pretzel noise. For the WCT matting algorithm, there was even a small increase in the number of consistent alpha masks as the pretzel noise level increased.

Fig. 9 shows the results of a combined comparison of different matting algorithms at different assessment levels. From this, it can be obtained that the Deep matting algorithm achieves the best results in the hierarchy of rigorous evaluation (Fig. 9a). The KNN matting algorithm has the best stability although it achieves the worst evaluation results. The amount of consistent alpha masking remained virtually unchanged across the four levels of pretzel noise. In the intermediate level of evaluation (Fig. 9b), all the matting algorithms were largely able to remain stable, with the LB matting algorithm showing the greatest fluctuations. The LB matting algorithm was able to achieve the best performance in the evaluation of the lax grade. From the evaluation results presented in Fig. 9, it can be concluded that pretzel noise has the greatest impact on the KNN matting algorithm, whereas the LB matting algorithm and deep matting algorithm are better adapted to the pretzel noise. The three matting algorithm can remain essentially stable across the three evaluation levels.

To better assess the effect of pretzel noise on the matting algorithm, we used MAD and RMSE to evaluate the alpha mask estimated in the pretzel noise images. Fig. 10 shows the results obtained by the matting algorithm in the quantitative assessment



Fig. 3 Estimated alpha mask in a Gaussian blur image. From left to right:

a Deep

b KNN

c LB

d Three

e WCT. The first row is the alpha mask estimated on the original image. The second to fourth rows show the alpha mask estimated on four different levels of Gaussian blur. The last four rows show the pixels in the noise alpha mask that are consistent with the original alpha mask (in the unknown area)

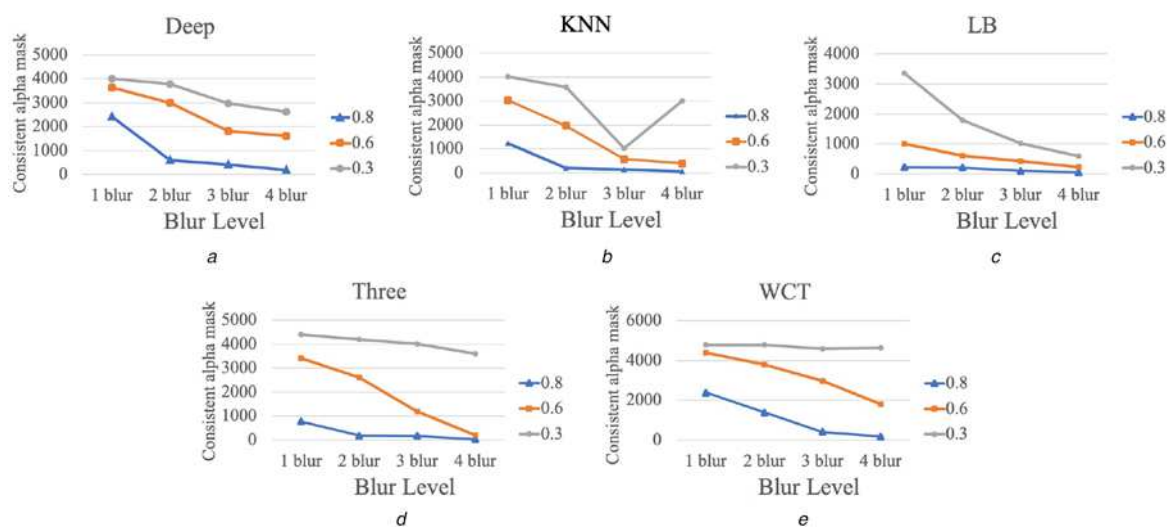


Fig. 4 Consistent evaluation of different matting algorithms in four levels of Gaussian blur

a Deep

b KNN

c LB

d Three

e WCT

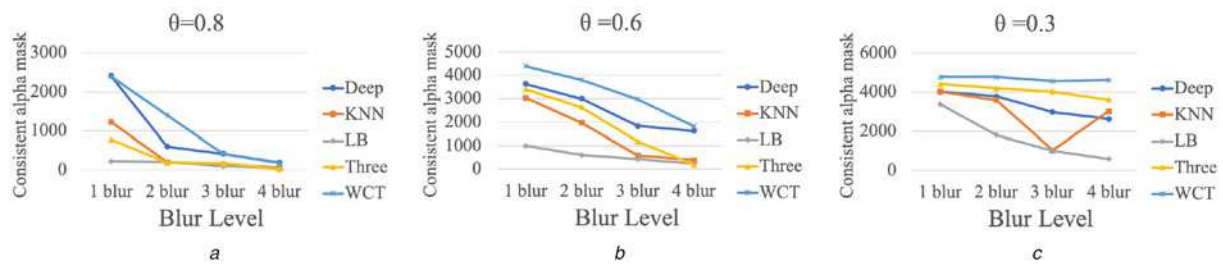


Fig. 5 Results of alpha mask consistency assessment in Gaussian blur. (a)–(c) The number of images that are consistent across the defined thresholds on the dataset by different matting algorithms

a 0.8
b 0.6
c 0.3

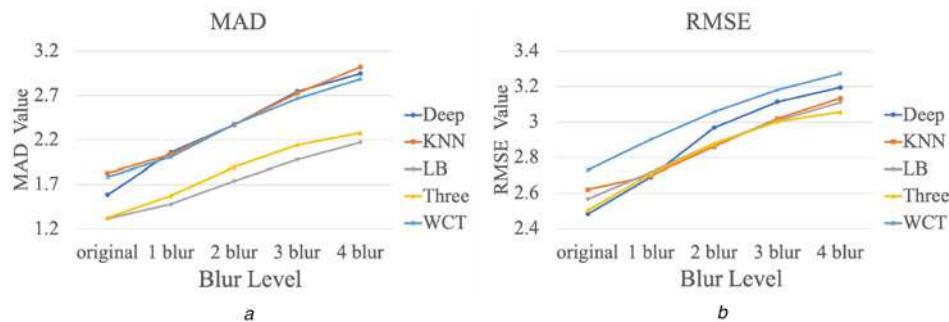


Fig. 6 Results of MAD and RMSE assessment in Gaussian blur

a MAD
b RMSE

of MAD and RMSE. In the MAD assessment results (Fig. 10b), the LB matting algorithm and the three matting algorithm were very close to the quantitative assessment results. These two matting algorithms give the best results in the quantitative assessment of MAD, whereas KNN matting algorithm and WCT matting algorithm performed poorly in the quantitative assessment of MAD. The deep matting algorithm and the three matting algorithm were more closely evaluated in the RMSE quantitative evaluation (Fig. 10a) and were the two matting algorithms with the best results in the RMSE quantitative evaluation. The WCT matting algorithm and the KNN matting algorithm continued to perform poorly in the quantitative evaluation of RMSE at different levels of pretzel noise. As the noise level of the pretzel noise increases, its RMSE quantitative assessment results become worse.

Fig. 11 shows the alpha mask estimated on the combined noise image of Gaussian blur and pretzel noise. The first row shows the alpha mask estimated on the original image. Rows two through four show the estimated alpha masking at four levels of combined noise. The last four rows show the consistent pixels of the combined noise alpha mask and the original alpha mask in unknown areas. The effect of combined noise on the matting algorithm is evident from the magnified region of Fig. 11. The effect on the estimated mixed pixels can be clearly observed in the combined noise blurred image at the fourth level. Combined noise has a more pronounced effect on the deep matting algorithm, KNN matting algorithm and LB matting algorithm. From the magnified region of similar pixels, the estimated combined noise alpha mask and the original alpha mask do not maintain a good consistency of pixels in the unknown region. The combined noise has less effect on the three matting algorithm and the WCT matting algorithm, and the estimated unknown pixels have better agreement at different levels of combined noise.

Fig. 12 shows the results of the specific evaluation of the matting algorithm for each test at different levels of combined noise. The effect of combined noise on the matting algorithm is more pronounced compared to Gaussian noise and pretzel noise. As the

noise level of the combined noise increases, the performance of the matting algorithm shows a significant decrease. From Fig. 12, it can be obtained that for the three matting algorithm and the WCT matting algorithm, the combined noise has less effect. The two algorithms are the least volatile under relaxed evaluation conditions. And for the deep matting algorithm, KNN matting algorithm, LB matting algorithm, the performance decreases significantly as the noise level increases.

Fig. 13 shows the results of a comprehensive comparison of different matting algorithms at different assessment levels. Under the more stringent evaluation conditions (Fig. 13a), the WCT matting algorithm can achieve the best evaluation results. The performance of the matting algorithms all showed significant degradation when the noise had a large impact on the original image. For the fourth level of combined noise, the results of the various matting algorithms are relatively close, and the WCT matting algorithm gives the best results for all three levels of evaluation (Table 3). And the LB matting algorithm did not perform well in the three levels of assessment. The WCT matting algorithm has the best stability in the evaluation of the lenient grades (Fig. 13c).

To better evaluate the effect of combined noise on the matting algorithm, we used MAD and RMSE to evaluate the alpha mask estimated in the combined noise image. Fig. 14 shows the quantitative evaluation of the combined noise alpha masking. In the MAD quantitative evaluation (Fig. 14a), the LB matting algorithm and the three matting algorithm showed the best results. In the quantitative evaluation of RMSE (Fig. 14b), the results of the individual matting algorithms were more similar.

Based on the previous analysis and the results we present, it can be obtained that the tested matting algorithms are all capable of estimating alpha masking on the noise image. The experimental results reflect that the matting algorithms are all sensitive to noise, but to different noises to different degrees. On the whole, Gaussian blurring has a greater effect on the matting algorithm compared to the pretzel noise. The estimated number of consistent

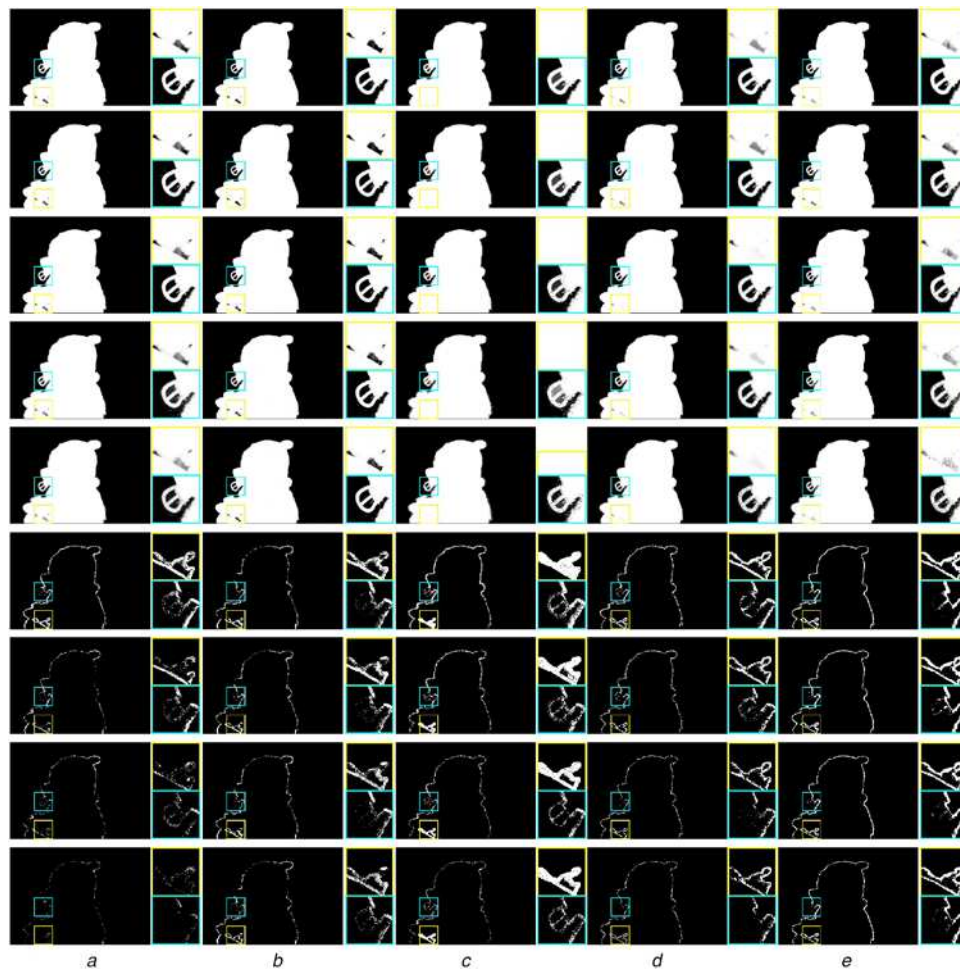


Fig. 7 Estimated alpha mask in a pretzel noise image. From left to right:

a Deep

b KNN

c LB

d Three

e WCT. The first row is the alpha mask estimated on the original image. The second to fourth rows show the alpha mask estimated on four different levels of pretzel noise. The last four rows show the pixels in the noise alpha mask that are consistent with the original alpha mask (in the unknown area)

Table 2 RMSE and MAD evaluation results (Gaussian blur)

	original		0.01		0.05		0.10		0.15	
	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD
deep	2.4834	1.5838	2.5123	1.6085	2.5178	1.5695	2.4903	1.5117	2.5203	1.6097
KNN	2.6205	1.8256	2.6365	1.8642	2.6499	1.8716	2.6710	1.9136	2.6816	1.9339
LB	2.5683	1.3154	2.5712	1.3291	2.5733	1.3399	2.5764	1.3559	2.5803	1.3654
three	2.5044	1.3253	2.5359	1.3522	2.5225	1.3352	2.5310	1.3473	2.5365	1.3421
WCT	2.7302	1.7838	2.7324	1.8172	2.7529	1.8375	2.7753	1.8595	2.7853	1.8619

Bold indicates the best performance.

alpha masks in Gaussian blurred images decreases rapidly with increasing noise levels. The number of consistent alpha masks decreased significantly as the assessment levels became progressively more stringent. The results of the quantitative evaluation showed that the performance of each matting algorithm is degraded to varying degrees in the presence of noise. To improve the robustness of the matting algorithm to noise, its ability to capture image textures to better capture adjacent pixels can be enhanced. For more severe noise, the boundaries of the foreground objects become blurred, which can enhance the ability of the matting algorithm to deal with weak boundaries to deal with the problem. The main objective of the algorithm in this paper is

to analyse the robustness of the matting algorithm under different noise conditions, while most of the evaluation methods are to evaluate the raw images without noise.

Discuss: According to the experimental results, the performance of the WCT matting algorithm is more robust and the performance of the KNN matting algorithm is most unstable in Gaussian fuzzy noise images. The performance of the WCT matting algorithm is more robust and the performance of the KNN matting algorithm is the most unstable in pretzel noise images. In the combined noise images, the performance of the WCT matting algorithm is more robust and the performance of the LB matting algorithm is the least stable. In practice, the WCT matting algorithm has better

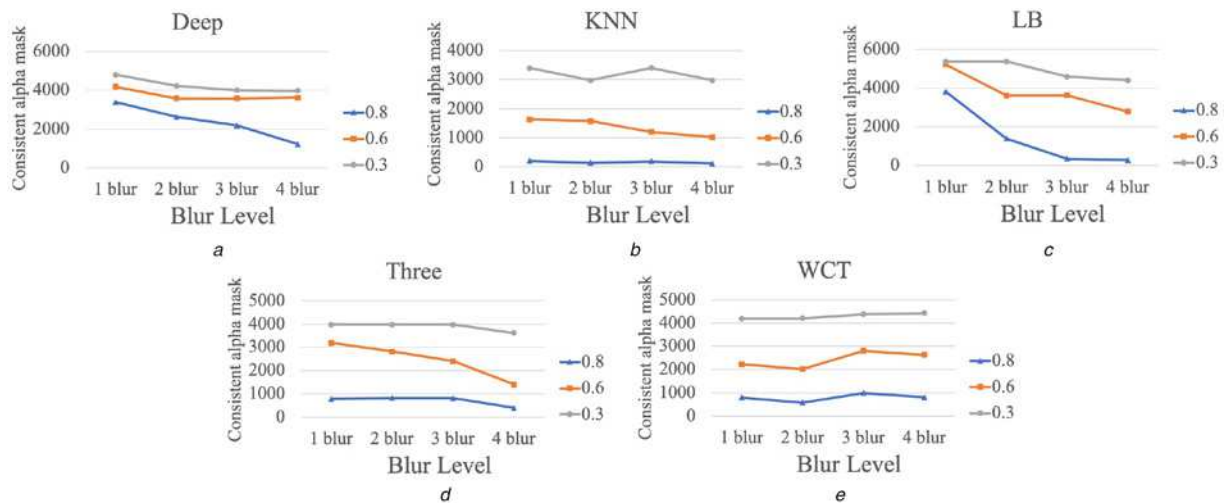


Fig. 8 Consistent evaluation of different matting algorithms in four levels of pretzel noise

a Deep
b KNN
c LB
d Three
e WCT

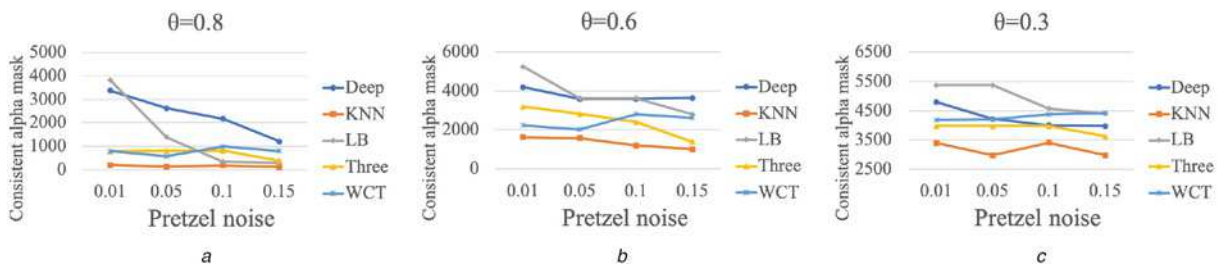


Fig. 9 Results of alpha mask consistency assessment in pretzel noise. (a)–(c) The number of images that are consistent across the defined thresholds on the dataset by different matting algorithms

a 0.8
b 0.6
c 0.3

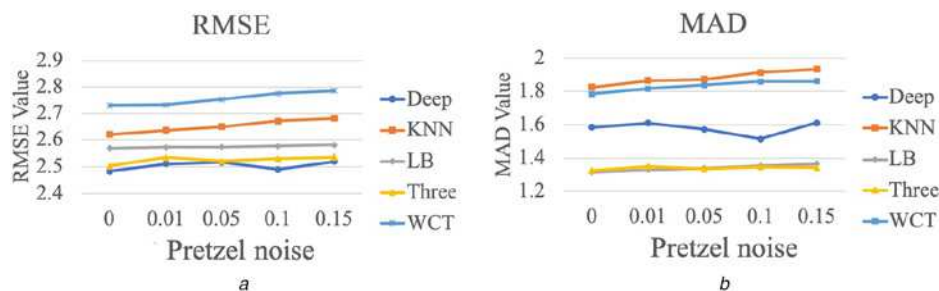


Fig. 10 Results of MAD and RMSE assessment in pretzel noise

a RMSE
b MAD

stability. Although the performance of the matting algorithm has been continuously improved as the technology continues to evolve. Among the MAD and RMSE evaluation indicators, the recently proposed matting algorithm has better evaluation results. However, these highly rated matting algorithms require more conditions for use, especially deep learning-based matting algorithms, which require large amounts of training data. It is possible to choose different matting algorithms for different applications, and there is no need to be obsessed with the latest

matting algorithm. We can choose the most appropriate matting algorithm for different usage environments. For example, matting algorithms with better resistance to noise will yield better results when extracting foreground objects from rain fog images or low-light images. In order to improve the robustness of the matting algorithm, methods such as image de-fogging, brightness enhancement, and image quality enhancement can be fused with the matting algorithm. This allows for more competitive results when dealing with poor quality images.

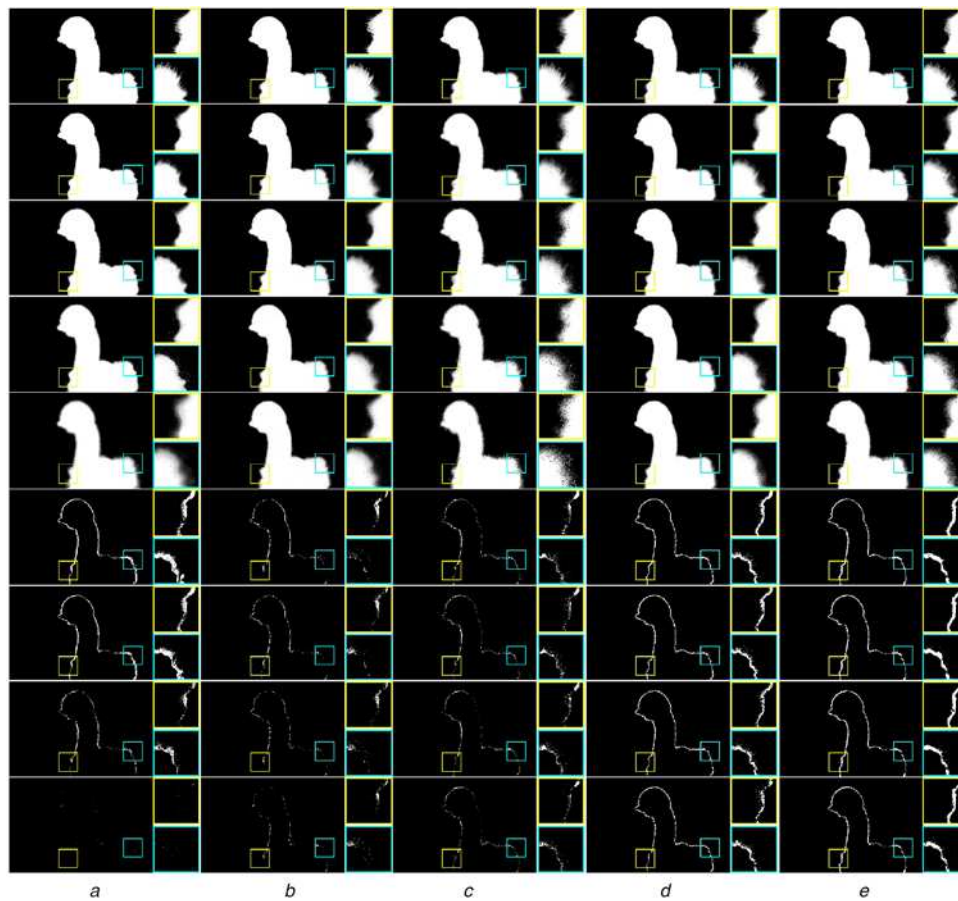


Fig. 11 Estimated alpha mask in combined noise image. From left to right:

a Deep
b KNN
c LB
d Three
e WCT

The first row is the alpha mask estimated on the original image. The second to fourth rows show the alpha mask estimated on four different levels of combined noise. The last four rows show the pixels in the noise alpha mask that are consistent with the original alpha mask (in the unknown area)

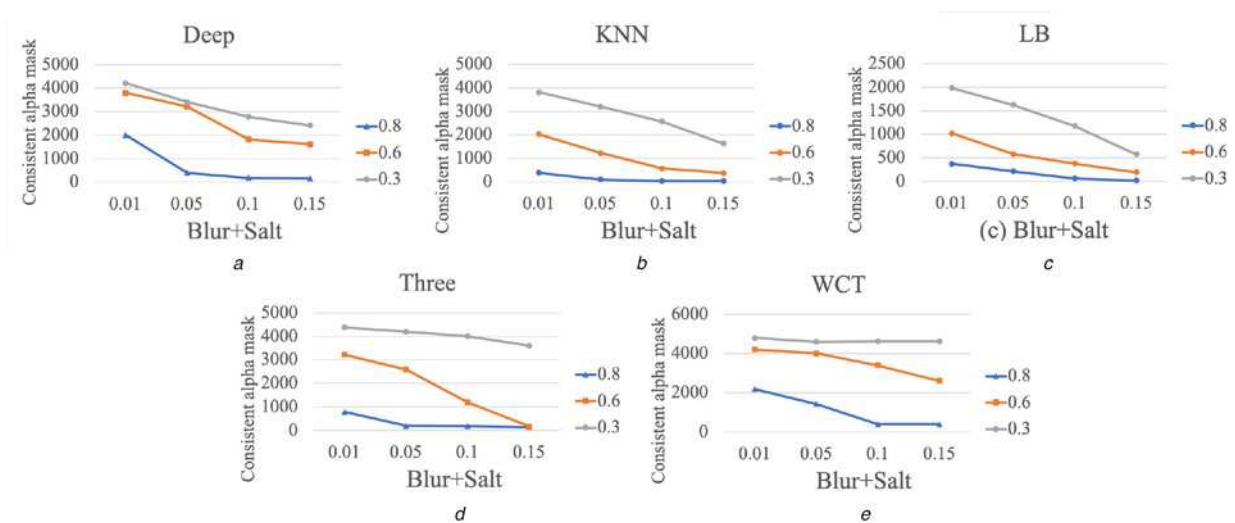


Fig. 12 Consistent evaluation of different matting algorithms in four levels of combined noise

a Deep
b KNN
c LB
d Three
e WCT

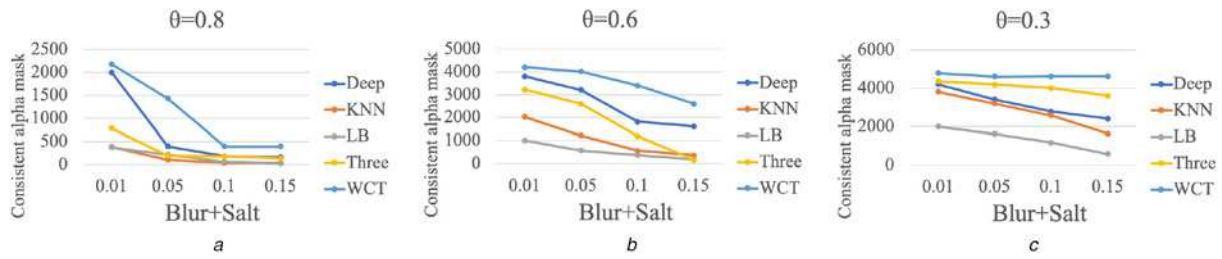


Fig. 13 Results of alpha mask consistency assessment in combined noise. (a)–(c) The number of images that are consistent across the defined thresholds on the dataset by different matting algorithms

a 0.8
b 0.6
c 0.3

Table 3 RMSE and MAD evaluation results (Gaussian blur + pretzel noise)

	original		0.01		0.05		0.10		0.15	
	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD
deep	2.4834	1.5838	2.6649	1.9359	2.9797	2.5790	3.1127	2.6391	3.1895	2.9170
KNN	2.6205	1.8256	2.7459	2.0690	2.9361	2.4529	3.0945	2.8044	3.2126	3.1395
LB	2.5683	1.3154	2.7191	1.4911	2.8800	1.7652	3.0119	2.0149	3.1145	2.2186
three	2.5044	1.3253	2.6931	1.5681	2.8731	1.8273	2.9879	2.0981	3.0295	2.2714
WCT	2.7302	1.7838	2.8914	2.0659	3.0517	2.4559	3.1412	2.7211	3.2054	2.9078

Bold indicates the best performance

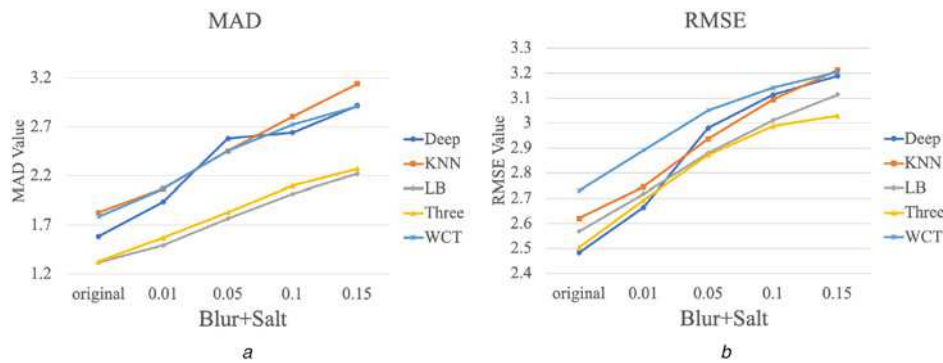


Fig. 14 Results of MAD and RMSE assessment in combined noise

a MAD
b RMSE

5 Conclusion

In this paper, we use a Gaussian–Hermite moment that combines gradient amplitude and gradient direction to evaluate the consistent alpha mask to determine the robustness of the matting algorithm. When varying degrees of noise are applied to the original image, the matting algorithm estimates that the noise alpha mask and the original alpha mask are said to be the same if they have more consistent pixels. We used different levels of consistency assessment to determine the degree of consistency of alpha masking. We use Gaussian blur, pretzel noise, and combined noise to destroy the image. We divided the alpha mask consistency into three levels: $\theta = \{0.3, 0.6, 0.8\}$. This corresponds to a relaxed grade ($\theta = 0.3$), an intermediate grade ($\theta = 0.6$), and a strict grade ($\theta = 0.8$), respectively. The full reference assessment approach we used when assessing the consistency of alpha masking. The results of the full reference evaluation are more realistic. Assessment methods based on deep learning are now developing rapidly. In future work, we are prepared to use deep neural networks to evaluate the quality of alpha masks. We are prepared to implement stepwise no-reference evaluation of the matting algorithm.

6 Acknowledgments

This research was supported by the National Natural Science Foundation of China (61772319, 61773244, 61976125, 61976124), Shandong Natural Science Foundation of China (ZR2017MF049) and Yantai Key Research and Development Plan (2019XDHZ081).

7 References

- [1] Kang, T.-K., Choi, I.-H., Lim, M.-T.: ‘MDGHM-SURF: a robust local image descriptor based on modified discrete Gaussian–Hermite moment’, *Pattern Recognit.*, 2015, **48**, (3), pp. 670–684
- [2] Khalid, H., Yasmeen, K., Walid, K., *et al.*: ‘Robust image hashing using exact Gaussian–Hermite moments’, *IET Image Process.*, 2018, **12**, (12), pp. 2178–2185
- [3] Wang, T., Zhang, L., Jia, H., *et al.*: ‘Multiscale contrast similarity deviation: an effective and efficient index for perceptual image quality assessment’, *Signal Process., Image Commun.*, 2016, **45**, pp. 1–9
- [4] Egiazarian, K., Ponomarenko, M., Lukin, V., *et al.*: ‘Statistical evaluation of visual quality metrics for image denoising’. Int. Conf. on Acoustics, Speech and Signal Processing, Calgary, Alberta, Canada, 2017, pp. 6752–6756

- [5] Kim, J., Lee, S.: 'Fully deep blind image quality predictor', *IEEE J. Sel. Top. Signal. Process.*, 2017, **11**, (1), pp. 206–220
- [6] Ghadiyaram, D., Bovik, A.C.: 'Massive online crowdsourced study of subjective and objective picture quality', *IEEE Trans. Image Process.*, 2015, **25**, (1), pp. 372–387
- [7] Jiang, C., Chen, Q., Yang, X.C.: 'Review on full reference image quality assessment algorithms', *Appl. Res. Comput.*, 2014, **31**, (1), pp. 13–22
- [8] Lukin, V.V., Ponomarenko, N.N., Ieremeiev, O.I., et al.: 'Combining full-reference image visual quality metrics by neural network', *Human Vis. Electron. Imaging XX*, 2015, **9394**, p. 93940K
- [9] Zerman, E., Valenzise, G., Dufaux, F.: 'An extensive performance evaluation of full-reference HDR image quality metrics', *Quality User Experience*, 2017, **2**, (1), p. 5
- [10] Zhang, L., Gu, Z., Liu, X., et al.: 'Training quality-aware filters for no-reference image quality assessment', *IEEE Multimedia*, 2014, **21**, (4), pp. 67–75
- [11] Li, Q., Fang, Y., Lin, W., et al.: 'Gradient-weighted structural similarity for image quality assessments'. IEEE Int. Symp. on Circuits and Systems, Lisbon, Portugal, 2015, pp. 2165–2168
- [12] Xue, W., Zhang, L., Mou, X., et al.: 'Gradient magnitude similarity deviation: a highly efficient perceptual image quality index', *IEEE Trans. Image Process.*, 2014, **23**, (2), pp. 684–695
- [13] Flusser, J., Farokhi, S., Th, H.C., et al.: 'Recognition of images degraded by Gaussian blur', *IEEE Trans. Image Process.*, 2015, **25**, (2), pp. 790–806
- [14] Zhou, W., Bovik, A.C.: 'Mean squared error: love it or leave it? A new look at signal fidelity measures', *IEEE Signal Process. Mag.*, 2009, **26**, (1), pp. 98–117
- [15] Levin, A., Ravacha, A., Lischinski, D.: 'Spectral matting'. IEEE transactions on pattern analysis and machine intelligence, 2007, vol. 30, (10), pp. 1699–1712
- [16] Liu, L., Hua, Y., Zhao, Q., et al.: 'Blind image quality assessment by relative gradient statistics and adaBoosting neural network', *Signal Process., Image Commun.*, 2016, **40**, pp. 1–15
- [17] Karacan, L., Erdem, A., Erdem, E.: 'Image matting with KL-divergence based sparse sampling'. IEEE Int. Conf. on Computer Vision, Santiago, Chile, 2015, pp. 424–432
- [18] Cho, D., Kim, S., Tai, Y. W., et al.: 'Automatic trimap generation and consistent matting for light-field images', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2016, **39**, (8), pp. 1504–1517
- [19] Wu, H., Li, Y., Miao, Z., et al.: 'A new sampling algorithm for high-quality image matting', *J. Vis. Commun. Image Represent.*, 2016, **38**, (C), pp. 573–581
- [20] Chuang, Y.-Y., Curless, B., Salesin, D.H., et al.: 'A Bayesian approach to digital matting'. IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, HI, USA, 2001, pp. 264–271
- [21] Wang, J., Cohen, M.F.: 'Optimized color sampling for robust matting'. Computer Vision and Pattern Recognition, Minneapolis, MN, 2007, pp. 1–8
- [22] Gastal, E.S., Oliveira, M.M.: 'Shared sampling for real-time alpha matting', *Comput. Graph. Forum*, 2010, **29**, (2), pp. 575–584
- [23] He, K., Rhemann, C., Rother, C., et al.: 'A global sampling method for alpha matting'. Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 2011, pp. 2049–2056
- [24] Shahrian, E., Rajan, D., Price, B., et al.: 'Improving image matting using comprehensive sampling sets'. Computer Vision and Pattern Recognition, Portland, OR, USA, 2013, pp. 636–643
- [25] Sun, J., Jia, J., Tang, C.-K., et al.: 'Poisson matting', *ACM Trans. Graph.*, 2004, **23**, (3), pp. 315–321
- [26] Levin, A., Lischinski, D., Weiss, Y.: 'A closed-form solution to natural image matting', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2007, **30**, (2), pp. 228–242
- [27] He, K., Sun, J., Tang, X.: 'Fast matting using large kernel matting Laplacian matrices'. Computer Vision and Pattern Recognition, San Francisco, CA, USA, 2010, pp. 2165–2172
- [28] Lee, P., Wu, Y.: 'Nonlocal matting'. Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 2011, pp. 2193–2200
- [29] Chen, Q., Li, D., Tang, C.-K.: 'KNN matting', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, **35**, (9), pp. 2175–2188
- [30] Aksoy, Y., Ozan Aydin, T., Pollefeys, M.: 'Designing effective inter-pixel information flow for natural image matting'. Computer Vision and Pattern Recognition, Honolulu, Hawaii, 2017, pp. 29–37
- [31] Cho, D., Tai, Y.-W., Kweon, I.: 'Natural image matting using deep convolutional neural networks'. European Conf. on Computer Vision, Amsterdam, the Netherlands, 2016, pp. 626–643
- [32] Xu, N., Price, B., Cohen, S., et al.: 'Deep image matting'. Computer Vision and Pattern Recognition, Honolulu, Hawaii, 2017, pp. 2970–2979
- [33] Lutz, S., Amliantitis, K., Smolic, A.J.A.P.A.: 'AlphaGAN: generative adversarial networks for natural image matting', arXiv preprint arXiv:1807.10088, 2018
- [34] Wee, C.Y., Paramesran, R., Mukundan, R., et al.: 'Image quality assessment by discrete orthogonal moments', *Comput. Vis. Pattern Recognit.*, 2010, **43**, (12), pp. 4055–4068
- [35] Sheikh, H.R., Bovik, A.C.: 'An information fidelity criterion for image quality assessment using natural scene statistics', *IEEE Trans. Image Process.*, 2005, **14**, (12), pp. 2117–2128
- [36] Wei, Z., Borji, A., Zhou, W., et al.: 'The application of visual saliency models in objective image quality assessment: a statistical evaluation', *IEEE Trans. Neural Netw. Learn. Syst.*, 2016, **27**, (6), pp. 1266–1278
- [37] Zhang, L., Zhang, L., Mou, X., et al.: 'FSIM: a feature similarity Index for image quality assessment', *IEEE Trans. Image Process. A Publication IEEE Signal Process. Soc.*, 2011, **20**, (8), p. 2378
- [38] Zhang, X., Feng, X., Wang, W., et al.: 'Edge strength similarity for image quality assessment', *IEEE Signal Process. Lett.*, 2013, **20**, (4), pp. 319–322
- [39] Lee, D., Lim, S.: 'Improved structural similarity metric for the visible quality measurement of images', *J. Electron. Imaging*, 2016, **25**, (6), p. 063015
- [40] Gupta, V., Chaurasia, V., Shandilya, M.: 'Random-valued impulse noise removal using adaptive dual threshold median filter', *J. Vis. Commun. Image Represent.*, 2015, **26**, pp. 296–304
- [41] Wang, F.-P., Shui, P.-L.: 'Noise-robust color edge detector using gradient matrix and anisotropic Gaussian directional derivative matrix', *Comput. Vis. Pattern Recognit.*, 2016, **52**, pp. 346–357
- [42] Shi, Z., Zhang, J., Cao, Q., et al.: 'Full-reference image quality assessment based on image segmentation with edge feature', *Signal Process.*, 2018, **145**, pp. 99–105
- [43] Shi, Z., Chen, K., Pang, K., et al.: 'A perceptual image quality index based on global and double-random window similarity', *Digit. Signal Process.*, 2017, **60**, pp. 277–286
- [44] Wn, Z., Li, Q.: 'Information content weighting for perceptual image quality assessment', *IEEE Trans. Image Process. A Publication IEEE Signal Process. Soc.*, 2011, **20**, (5), pp. 1185–1198
- [45] Indu, S., Ramesh, C.: 'A noise fading technique for images highly corrupted with impulse noise'. Int. Conf. on Computing: Theory and Applications, Kolkata, India, 2007, pp. 627–632
- [46] Ni, Z., Ma, L., Zeng, H., et al.: 'Gradient direction for screen content image quality assessment', *IEEE Signal Process. Lett.*, 2016, **23**, (10), pp. 1394–1398
- [47] Bae, S.-H., Kim, M.: 'A novel image quality assessment with globally and locally consistent visual quality perception', *IEEE Trans. Image Process.*, 2016, **25**, (5), pp. 2392–2406
- [48] Brekhna, B., Mahmood, A., Zhou, Y., et al.: 'Robustness analysis of superpixel algorithms to image blur, additive Gaussian noise, and impulse noise', *J. Electron. Imaging*, 2017, **26**, (6), p. 1
- [49] Chen, D.-J., Chen, H.-T., Chang, L.-W.: 'Fast defocus map estimation'. IEEE Int. Conf. on Image Processing, Phoenix, AZ, USA, 2016, pp. 3962–3966