Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 16, No. 1, 2025

Depth-aware Image-Space Fogging Algorithm for Background Privacy Protection using ZoeDepth Framework

Xiyin Liang[1,2,*], Meng Wang[1,2], Peirong Pan[1,2]

[1] College of Physics and Electronic Engineering, Northwest Normal University, Peili Street, Lanzhou 730070, Gansu, China

[2] Engineering Research Center, Gansu Province for Intelligent Information Technology and Application, Peili Street, Lanzhou, 730070, Gansu, China

[*] **Corresponding author:** Xiyin Liang (Email: silver@nwnu.edu.cn)

Abstract: Fogging algorithms are typically employed in fog dataset construction, natural scene generation for games, and fog scene rendering. This paper proposes a novel depth-aware image-space fogging algorithm for background privacy protection, leveraging the ZoeDepth depth estimation framework and atmospheric scattering model. The algorithm first estimates depth maps from single images using ZoeDepth, enabling segmentation into foreground and background regions. Subsequently, by integrating the atmospheric scattering model and an improved dark channel prior algorithm, the algorithm applies fogging processing to the background regions, transitioning naturally from near to far without retaining identifiable information. Evaluation metrics, including Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Information Entropy, demonstrate that the proposed approach preserves details and structural information in the foreground while effectively blurring background information. This method offers a practical solution for image privacy protection in diverse domains such as security, social media, and military applications, where background privacy is paramount.

Keywords: ZoeDepth model; Depth Estimation; Fogging Algorithm; Background Privacy Protection; Depth-aware Image-Space.

1. Introduction

The rapid development of technology and the increasing supply and demand of visual information in various sectors of society have led to serious concerns about sensitive information leakage. Digital privacy protection of visual information, represented by visual information, has become a hot ethical issue and a profound concern for people in response to the proliferation of information. Currently, commonly used visual privacy protection methods represented by image information mainly focus on information masking and weakening measures, such as pixelation, blurring, replacement, and fog processing, targeting the main subject of the image. Zhou et al. applied pixelation-based privacy protection methods to live video streaming [1]. However, pixelation methods suffer from the drawback of losing detailed information, leading to misunderstandings of the protected objects or scenes. Salman et al. proposed a privacy protection operation based on the YOLOv6 algorithm, which protects individual privacy while detecting faces through encryption, masking, and blurring [2]. Li et al. addressed privacy protection issues when sharing photos on online social networks, proposing a facial privacy protection method based on blurring [3]. Khojaste et al. introduced a new model called GMFIM, aiming to hide personal identity by altering facial image features [4]. These foreground information masking or weakening methods have been effective in protecting explicit foreground privacy information. However, the backgrounds of images captured by visual capture devices often contain extremely rich sensitive information. For example, visual monitoring mechanisms such as sentry mode used in some autonomous vehicles can record vast amounts of parking and driving scene information. Electronic peepholes widely used in residential buildings can record neighbors' daily activities. The background privacy leakage caused by these visual capture devices without any privacy protection mechanisms has become an increasingly serious social and ethical issue.

Image background privacy protection refers to safeguarding the privacy of non-primary individuals in images, preventing these individuals from being publicly displayed or exploited without consent, thus avoiding the exposure or misuse of their identities and locations [5]. To address this issue, systems such as PUPPIES and AutoPri, proposed in articles [6,7], partition sensitive areas in photos to achieve privacy protection. Hasan et al. proposed a method for hiding sensitive information using image filters, which can effectively achieve image privacy protection [8]. Privacy protection algorithms proposed in articles [9,10] are used to obstruct the background information in classified images on social media. However, current background privacy protection faces two technical challenges: accurate recognition, judgment, and segmentation of image backgrounds based on image spatial perception, and weakening of background information with natural effects.

In previous studies, most research focused on dehazing methods and image restoration techniques [11-13]. In 2019, Sheng et al. proposed a depth-aware motion blur removal method by extending the Richardson-Lucy algorithm. This method combines depth images with a motion blur model, using depth information to fill in gaps and iteratively optimize both deblurring and depth recovery [14]. In 2022, Zhou et al. introduced the feedback spatial attention dehazing network (FSAD-Net), which, through a recursive structure and four core modules, effectively exploited the interrelationships between intermediate layer features, improving dehazing

performance. Experimental results showed that FSAD-Net outperforms existing state-of-the-art methods on multiple quantitative metrics and demonstrates superiority in real-world images [15]. Research on fogging processing, however, remains limited. Fogging is a type of image information attenuation process with a gradual natural transition. From 2013 to 2023, Researchers have primarily applied fogging processing to the construction of foggy image datasets and its application in the field of autonomous driving[16-23]. Additionally, the dataset provided by the articles[24-26], based on Convolutional Neural Networks, Nighttime Haze Rendering Module, and small-scale climate chamber equipment, serves as a dataset for the dehazing algorithm. Most synthetic foggy images are generated using uniform fogging, lacking the natural transition accompanying depth. While they serve as good experimental samples for dehazing algorithms, they do not meet the needs of background privacy protection. Simulating real-world natural fog effects, where fog density gradually increases with depth space, is a method of significantly weakening background information while maintaining image quality. We refer to this as "image-space fogging."

Deep learning is a branch of machine learning that automatically extracts features and makes predictions through multi-layer neural networks. Its application areas are extensive, including computer vision, natural language processing, autonomous driving, medical image analysis, and financial risk prediction. In recent studies, Li et al. proposed a reflectance-guided network, RADepthNet, which improves monocular depth estimation accuracy through reflectance guidance and boundary feature fusion [27]. Wang et al. proposed a robust blind watermarking algorithm combining stable interest points and deep learning networks, enhancing watermark extraction accuracy and robustness under mixed and geometric attacks [28].This paper proposes a concept and algorithm of image-space fogging, combining the ZoeDepth depth estimation model and the dark channel prior algorithm, to achieve image background privacy protection. The innovations of this paper include:

1). Utilizing the ZoeDepth monocular depth estimation model to obtain depth information from a single original image, eliminating the need for multiple images.

2). Comparing with uniform fogging, this paper segments the original image into foreground and background regions based on depth information, and applies fogging processing to the background region. This approach weakens the privacy information in the background region in a gradient manner along the depth direction while preserving the intrinsic information value of the image.

3). Integrating the ZoeDepth depth estimation model and the dark channel prior algorithm, a new image-space fogging algorithm is proposed.

The remaining chapters of this paper are arranged as follows. Section 2 primarily introduces the system framework of the image-space fogging algorithm. Section 3 presents the ZoeDepth model used for depth estimation. Section 4 mainly introduces the image-space fogging algorithm. Section 5 consists of experimental simulation and result analysis. Finally, Section 6 provides the conclusion. The numerical simulation environment for the data in this paper is used Intel (R) Core (TM) i7-8700 CPU @ 3.20 GHz, 16GB RAM, and the Windows 10 operating system.

2. System Framework

The system framework of the proposed image-space fogging algorithm is illustrated in Fig. 1. It mainly consists of four modules: the image input module, the depth estimation module, the fogging processing module, and the display terminal. The image input module is responsible for inputting original images captured in different scenes. The depth estimation module utilizes image information combined with the ZoeDepth depth estimation algorithm to generate depth maps corresponding to the input images. The fogging processing module includes the atmospheric scattering model and an improved dark channel prior algorithm. The atmospheric scattering model simulates fogging scenes in reality, combined with the depth map of the original image and the improved dark channel prior algorithm, to achieve fogging effects at different distances, making the fogging processing more realistic. The image transmission module is responsible for transmitting the fogged images to different application scenarios.

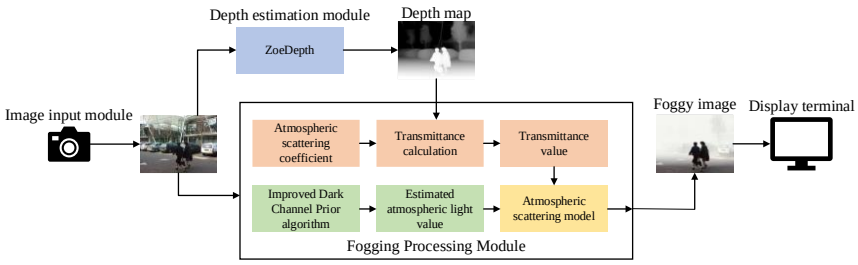

Figure 1. System block diagram.

Currently, there are various depth information parsing algorithms for indoor and outdoor scenes [29-33]. The method proposed in this paper for generating depth maps adopts monocular depth estimation technology, which combines neural networks with traditional computer vision methods. The basic principle is to utilize input monocular images to predict the depth distance information of each pixel through depth estimation algorithms, thereby constructing the structure of the scene. The atmospheric scattering model is a classical physical model used to simulate the phenomenon of light scattering in the atmosphere, thereby generating realistic haze effects. The model is based on the interaction of light with gases and aerosol particles in the atmosphere, resulting in the scattering and absorption of light. By using the atmospheric scattering model, it is possible to effectively simulate realistic haze effects in different scenes, providing important physical basis and credibility for fogging processing. Finally, the dark channel prior algorithm is used to estimate the atmospheric light value, and this value is then substituted into the mathematical formula of the atmospheric scattering model to obtain the final fogged image. The dark channel prior algorithm assumes that most areas in natural

scenes have low brightness values, indicating the presence of a haze-free dark channel. This algorithm utilizes this prior knowledge by estimating the dark channel through searching for the minimum value of pixels within a local window, thereby inferring the atmospheric light value of the image. By substituting the atmospheric light value and the transmittance obtained through depth information into the atmospheric scattering model formula, the input haze-free image can be transformed into a fogged image with background privacy information obscured.

3. Monocular depth vision generation model

ZoeDepth is an efficient and accurate depth estimation open-source project developed by the Image and Signal Laboratory [34]. The project combines a two-stage framework consisting of relative depth pre-training and metric depth fine-tuning, resulting in a model, ZoeD-M12-NK, that maintains measurement scale while achieving excellent generalization performance. This model is used for monocular depth estimation from single images, bridging the gap between relative and absolute depth estimation performance. It achieves outstanding generalization performance while maintaining measurement scale.

3.1. Overview of ZoeDepth method

The model framework is illustrated in Fig. 2, which includes a relative depth estimation network and an metric depth estimation network. The relative depth estimation network learns to extract depth difference information between adjacent pixels, while the metric depth estimation network directly predicts absolute depth values. With this framework, the model can transfer depth information from existing datasets to new target datasets, thus achieving zero-shot depth estimation.

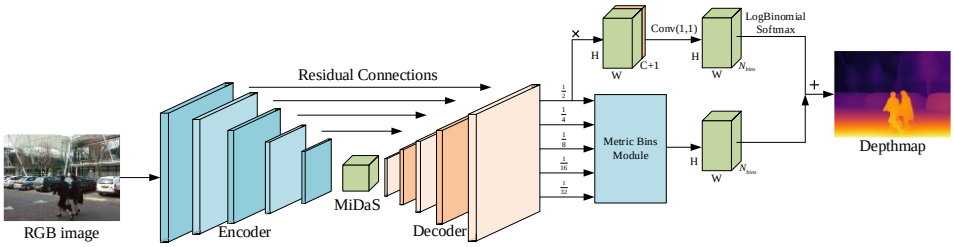

Figure 2. ZoeDepth model architecture.

In the first stage, the model utilizes a generic encoder-decoder architecture for pre-training on relative depth estimation. In the second stage, a lightweight head for metric depth estimation is added, and fine-tuning is performed using an metric depth dataset. Initially, a backbone consisting of an encoder-decoder is used for relative depth prediction. Then, a proposed metric bins module is attached to the decoder to obtain the metric depth prediction head. Absolute depth estimation is carried out by adding one or more heads, followed by end-to-end fine-tuning. LocalBins uses a standard encoder-decoder as the base model, supplemented by a module that takes the multi-scale features of the encoder-decoder as input to predict the center values of bins for each pixel's depth interval. The final pixel depth is obtained by a linear combination of the center values of bins, weighted by the probabilities derived from applying the softmax function.

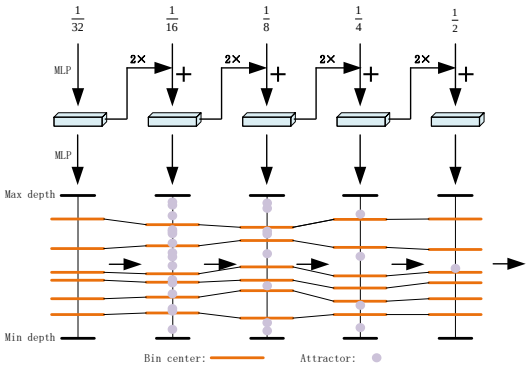

Figure 3. Metric Bins module.

The block diagram of the Metric Bins module is shown in Fig. 3. The module takes multi-scale features from the decoder of the MiDaS network as input and predicts the center of bins used for metric depth estimation. In this project, all bins for each pixel are directly predicted at the bottleneck layer, and then attractor layers are used on the decoder to progressively refine the bin intervals. As shown in the figure, there are $N_{\text{total}} = 64$ bins set for the depth intervals, and the decoder employs {16, 8, 4, 1} attractors. In the first decoder layer, an MLP takes the features at a pixel location as input and predicts the attractor point $\{a_k; k = 1, \ldots, na\}$ for that pixel position, adjusting the center c_i' of the bin accordingly:

$$\begin{cases} c_i' = c_i + \Delta c_i \\ \Delta c_i = \sum_{k=1}^{n_a} \dfrac{a_k - c_i}{1 + \alpha \left| a_k - c_i \right|^{\gamma}} \end{cases} \tag{1}$$

where α and γ represent the attractor strengths, named as inverse attractors. In the depth prediction task, each pixel needs to estimate its depth value, which may lie within a continuous depth interval. To obtain the final metric depth prediction, the project adopts a probability prediction method based on the binomial distribution. The model connects the relative depth prediction with the decoder features and predicts a two-channel output from the decoder features. In this way, the model can more accurately predict the probability distribution of each bin within the depth interval, thereby more accurately estimating the depth value of each pixel. To obtain the probability of the k-th bin center:

$$p(k; N, q) = \binom{N}{k} q^k (1-q)^{N-k} \tag{2}$$

After obtaining the probability distribution of each bin, the softmax function is used to normalize them to obtain the final probability values:

$$\text{softmax}\left(\left\{ \log\left(p_k \right) / t \right\}_{k=1}^{N} \right) \tag{3}$$

where t represents the temperature coefficient. Finally, the final depth of the pixel is obtained through Eq. 4:

$$d(i) = \sum_{k=1}^{N_{total}} p_i(k) c_i(k) \tag{4}$$

3.2. Experimental results

Four images selected from the UCID.V2 dataset are named Street, Plane, Car, and Sea, with all image sizes set to 640×480. The depth maps corresponding to each original image are shown in Fig. 4. From the images, it can be observed that the depth maps of different images exhibit different characteristics in terms of object positions, shapes, and structures.

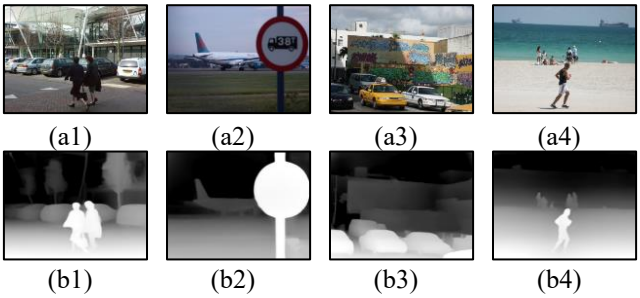

Figure 4. Different images and their corresponding depth maps. (a1) Original image of Street. (a2) Original image of Plane. (a3) Original image of Car. (a4) Original image of Sea. (b1) Depth map corresponding to Street. (b2) Depth map corresponding to Plane. (b3) Depth map corresponding to Car. (b4) Depth map corresponding to Sea.

4. Image-Space Fogging algorithm

4.1. Image-Space Fogging algorithm framework

The framework of the image-space fogging algorithm proposed in this paper is illustrated in Fig. 5. Firstly, the basic theoretical model of the algorithm is derived based on the mathematical model of atmospheric scattering. Then, the estimation of atmospheric light is derived according to the mathematical model of the original dark channel prior algorithm, and the transmittance is calculated based on depth information [35]. The original image is segmented into foreground and background regions by setting a threshold for depth information. By substituting the values of transmittance, atmospheric light, and the input image into the atmospheric scattering model, fogging processing based on depth information can be applied to the target image, achieving background fogging.

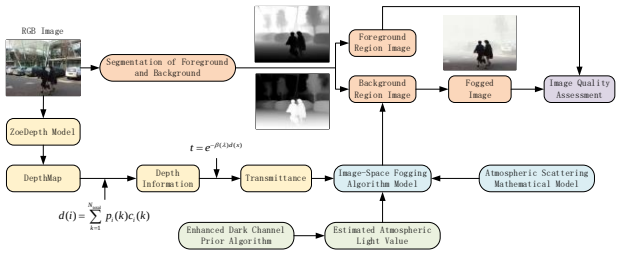

Figure 5. Algorithm flowchart.

4.2. Atmospheric scattering model

The atmospheric scattering model is a mathematical model used to describe the propagation of light and the distribution of illumination in the atmosphere[36]. The model primarily simulates the processes of scattering, absorption, and attenuation of sunlight in the atmosphere, as well as the effects of these processes on the intensity and color of illumination at the ground or observation points. In 1999, Narasimhan et al. established the atmospheric scattering model, as shown in Fig.6[37, 38]. The atmospheric scattering model is also commonly applied in game development and virtual reality applications to study how to construct models to achieve more realistic rendering of virtual fog scenes[39]. From the figure, it can be seen that the light received by the observation point mainly comes from two parts: one part is the reflected light from the target object to the light source, which arrives at the observation point after attenuation by particle propagation. The other part is the atmospheric ambient light formed by particle scattering during the propagation of light in the atmosphere. The mathematical expression of this model is shown in Eq. (5):

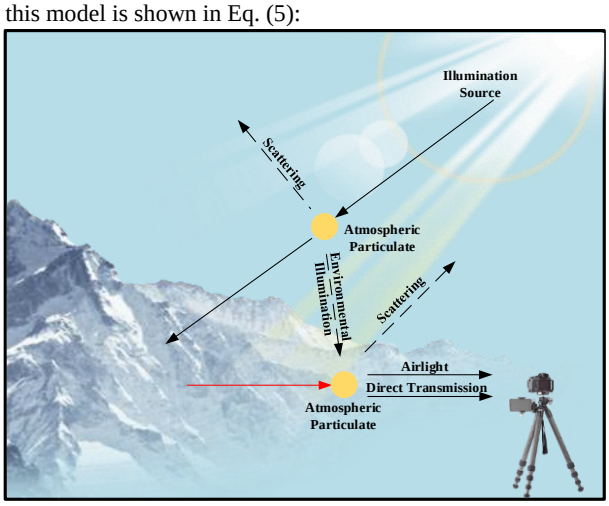

Figure 6. Atmospheric scattering imaging model.

$$\begin{cases} I(x) = D + A = J(x)t(x) + A_\infty(1 - t(x)) \\ D = A_\infty J(x)(1 - t(x)) = A_\infty(1 - t) \end{cases} \tag{5}$$

where D represents the incident light from the target object to the observation point, which is equal to the product of the reflected light from the target $J(x)$ and the transmittance $t(x)$; A represents the atmospheric ambient light; $I(x)$ represents the foggy image observed at the observation point; x represents the pixel coordinates of the image; A_∞ represents the atmospheric light value at infinity. The Eq.for calculating transmittance is shown in Eq. (6). $\beta(\gamma)$ is the scattering coefficient, usually taken as a constant $\beta = 3.912/R_m$; $d(x)$ represents the depth information in the depth map.

$$t = e^{-\beta(\lambda)d(x)} \tag{6}$$

4.3. Dark channel prior algorithm

The main goal of the dark channel prior algorithm is to estimate the atmospheric light intensity in Eq. (6). The "dark channel" refers to the channel within a local region of the image that has the lowest brightness value. According to the theory proposed by [35] in the dark channel prior algorithm, the mathematical expression of its dark channel is defined as:

$$J^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in (r,g,b)} J^c(\mathbf{y}) \right) \tag{7}$$

where x represents the pixel value of an image, R, G, and B represent the color channels in a color image, $J^{dark}(x)$ represents the dark channel value of the haze-free image, $\Omega(x)$ represents a local block centered at coordinate x,

c represents a three-dimensional matrix composed of the RGB channels, and $J^c(\mathbf{y})$ represents one color channel in a natural image. Each pixel value in the dark channel image represents the darkest pixel value in the corresponding region of the original image, reflecting the minimum light intensity in each region of the image. In atmospheric light estimation, the dark channel is used to find the sky region in the image, as the sky region usually has a higher brightness, which will appear as lower values in the dark channel image. Therefore, approximating the grayscale values of all pixels in the entire dark channel image to zero can be represented as:

$$J^{dark} \to 0 \tag{8}$$

First, according to the dark channel prior, the top 0.1% brightest pixels are selected from the dark channel image to avoid interference from natural image scenes. Then, from these brightest pixels, find the pixel with the median grayscale value and use it as the estimated atmospheric light pixel. The atmospheric light is represented by a vector R, G, B containing the values of the red, green, and blue channels of the atmospheric light. The index containing the position of the estimated atmospheric light pixel is used to determine the position of the atmospheric light in the original image. Based on the given dark channel image, the atmospheric light of the original image is estimated, and the color value of the estimated atmospheric light along with its position in the original image is returned.

4.4. Algorithm steps

The specific steps of the image space fogging algorithm are as follows:

Step 1: Calculate the transmission rate based on the depth map and the atmospheric scattering model. The specific implementation steps are shown in Algorithm 1

Algorithm 1: Transmittance Calculation
Input: Depth map; Atmospheric scattering coefficient;
Step 1: Import necessary libraries; Step 2: Calculate the distance from each pixel to the camera; Step 3: Define the maximum and minimum fog intensities; Step 4: Calculate fog intensity based on distance, ensuring that pixels farther away correspond to higher fog intensities and pixels closer correspond to lower fog intensities, while limiting fog intensity within a certain range; Step 5: Compute transmittance using the formula. Output: Transmittance of each corresponding pixel.

Step 2: Utilize the dark channel prior algorithm to estimate the atmospheric light value from the image. The specific implementation steps are shown in Algorithm 2:

Algorithm 2: Estimation of Atmospheric Light Using Dark Channel Prior Algorithm
Input: Dark channel image; Original color image;
Step 1: Define the fraction of the brightest pixels used for estimating atmospheric light as 1/1000 of the dark channel pixels. Step 2: Calculate the total number of pixels in the image. Step 3: Determine the number of brightest pixels in the dark channel. Step 4: Sort the pixels in the dark channel to find the brightest pixels and save their indices. Step 5: Sort the pixels in the original image and find the corresponding brightest pixels from the dark channel. Step 6: Select the median grayscale value from the brightest pixels in the original image and obtain the corresponding pixel position in the original image as the estimated atmospheric light. Output: Color vector of the estimated atmospheric light; Indices of pixels in the original image.

Step 3: Applying the atmospheric scattering model to fog the input image using the estimated atmospheric light and transmittance rate. Utilizing the Eq.of the atmospheric scattering model, the image space fogging algorithm is applied to each pixel. Different concentrations of fog are applied in the depth direction of the image in a gradient weakening manner. Finally, the fogged image is output, where the privacy regions have been blurred while the non-privacy regions remain clear, serving the purpose of background privacy protection.

5. Experimental results and performance evaluation

5.1. Experimental results

Four images, named Street, Plane, Car, and Sea from Fig.4, are selected for fogging using the proposed image space fogging algorithm. The simulation results are depicted in Fig. 7. Figs. 7(a1-a4) represent the original images, Figs. 7(b1-b4) represent the computed transmittance maps, and Figs. 7(c1-c4) represent the images after fogging. It can be observed that the background areas of the images after fogging have minimal recognizable information visually, indicating that the employed image space fogging algorithm successfully protects the sensitive information in the images. Furthermore, the foreground areas of the fogged images remain identical to the original images, demonstrating that the foreground regions remain undamaged during the fogging process. Additionally, when processing images under different scenes and lighting conditions, the algorithm preserves the quality and authenticity of the images, unaffected by scene changes. This indicates that the proposed image space fogging algorithm exhibits better universality and stability.

Figure 7. Simulation results of image Space Fogging algorithm. (a1) Original image of Street. (a2) Original image of Plane. (a3) Original image of Car. (a4) Original image of Sea. (b1) Transmittance map corresponding to Street. (b2) Transmittance map corresponding to Plane. (b3) Transmittance map corresponding to Car. (b4) Transmittance map corresponding to Sea. (c1) Fogged image corresponding to Street. (c2) Fogged image corresponding to Plane. (c3) Fogged image corresponding to Car. (c4) Fogged image corresponding to Sea.

5.2. Comparative analysis

Five images, named Corridor, Hallway, Grocery, Couch, and Boardroom, are selected from the NYU indoor dataset as sample images for comparative analysis. The images are resized to 640 × 480 pixels. The proposed image space fogging algorithm is compared with the domain-adaptation fogging algorithm proposed in reference [40]. The fogged images obtained using different algorithms are shown in Fig. 8. Visually, it can be observed that our method provides better differentiation between the foreground and background regions of the images. The images in the foreground areas exhibit nearly identical clarity, detail preservation, and color fidelity to the original images. Particularly in complex scenes with textured and detailed foreground areas, our method better preserves the details and structure of the images. The fog applied to the background areas makes objects in the background difficult to discern, thereby better serving the purpose of background privacy protection. In contrast, the fogged images generated using the domain-adaptation synthesized fogging method do not differentiate between the foreground and background regions, and the applied fog concentration does not effectively blur object information in the images.

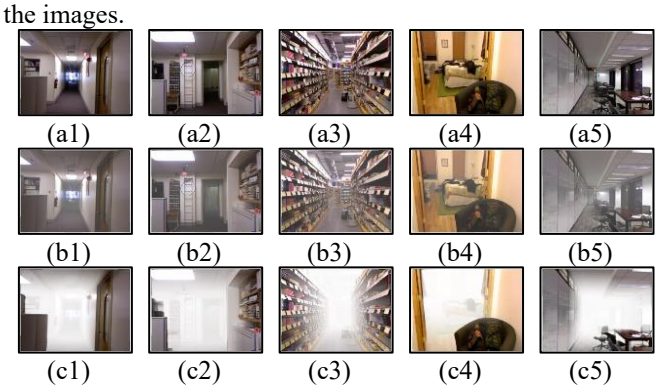

Figure 8. Contrast results. (a1) Original image of Corridor. (a2) Original image of Hallway. (a3) Original image of Grocery. (a4) Original image of Couch. (a5) Original image of Boardroom. (b1) -(b5) Simulated results of Domain Adaptation Synthetic Fogging Algorithm. (c1) -(c5) Simulated results of the proposed algorithm in this paper.

5.3. Quality evaluation

Four commonly used image quality assessment metrics [41], including Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Information Entropy, are employed to evaluate the quality of the fogged images. MSE is calculated according to Eq. (9), where R(i, j) represents the pixel of the original image (i, j), F(i, j) represents the pixel of the fogged image (i, j), and M×N represents the size of the image. PSNR is computed as shown in Eq. (10), where the relationship between MSE and image quality is inversely proportional, while PSNR is directly proportional to image quality. This means that the smaller the MSE, the larger the PSNR, indicating a lower deviation between the reference and evaluated images, and thus a better image quality.

The SSIM calculation formula is shown in Eq. (11), where ux is the mean value of x, uy is the mean value of y, is the variance of x, is the variance of y, and is the covariance of x and y, c1=(k1L)2 and c2=(k2L)2 are constants used to stabilize the division with weak denominators. L is the dynamic range of pixel values. And k1 = 0.01 and k2 = 0.03 are used to maintain stability. When two images are identical, the SSIM value is 1.

Information entropy is computed according to Eq. (12), where p(mi) represents the probability of a sample, and 2n represents the total number of samples. A higher information entropy indicates more complex and varied details and structures in the image. Conversely, a lower information entropy suggests that the pixel values in the evaluated image are more regular, presenting smoother and simpler features with smaller differences between adjacent pixels.

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|R(i,j)-F(i,j)|^2 \tag{9}$$

After obtaining the probability distribution of each bin, the softmax function is used to normalize them to obtain the final probability values:

$$PSNR = 10\lg\frac{255^2}{MSE} \tag{10}$$

where t represents the temperature coefficient. Finally, the final depth of the pixel is obtained through Eq. 4:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{11}$$

$$H(m) = -\sum_{i=0}^{2^n-1} p(m_i)\log_2 p(m_i) \tag{12}$$

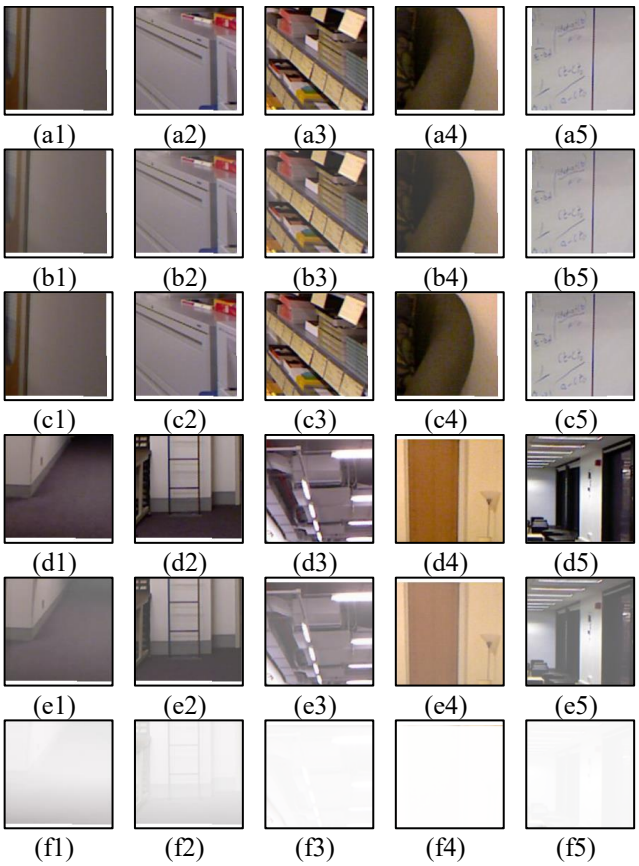

Figure 9. Contrast results of foreground and background areas. (a1) -(a5) Foreground areas of original images. (b1) -(b5) Domain Adaptive Synthetic Fogging Algorithm for foreground areas. (c1) -(c5) The proposed algorithm in this paper for foreground areas. (d1) -(d5) Background areas of original images. (e1) -(e5) Domain Adaptive Synthetic Fogging Algorithm for background areas. (f1) -(f5) The proposed algorithm in this paper for background areas.

The comparison of the image-space fogging algorithm proposed in this paper with the domain-adaptation fog synthesis algorithm before and after fogging is shown in Fig. 9 for subjective evaluation results, and in Table 1 for objective

evaluation results [37]. As seen, compared to the domain-adaptation fog synthesis algorithm, the proposed algorithm achieves an MSE value of 0 for the foreground region, while the background region's MSE value is significantly higher than that of the domain-adaptation fog synthesis algorithm. This indicates that the fogging process preserves the pixel values of the foreground region perfectly, maintaining the image information intact. In contrast, the background region experiences significant distortion, effectively protecting background privacy. The MSE values for the domain-adaptation fog synthesis algorithm suggest that the foreground region's image is affected by the fogging process and does not retain the information as well.

Table 1. Image quality evaluation results.

Metrics	Region	Algorithm	Corridor	Hallway
MSE	foreground area	ours	0	0
		others	143.1070	40.6752
	background area	ours	22047.4264	19403.8890
		others	1513.6301	587.3638
PSNR	foreground area	ours	Inf	Inf
		others	26.5742	32.0375
	background area	ours	4.6972	5.2519
		others	16.3306	20.4417
SSIM	foreground area	ours	1	1
		others	0.9869	0.9940
	background area	ours	0.5892	0.4996
		others	0.8750	0.9262
E	foreground area	original image	5.8181	6.1939
		ours	5.8181	6.1939
		others	5.5782	6.0661
	background area	original image	6.5689	6.5748
		ours	5.4375	5.1317
		others	5.9299	6.4087

Metrics	Region	Algorithm	Grocery	Couch	Boardroom
MSE	foreground area	ours	0	0	0
		others	216.0177	499.3907	13.3029
	background area	ours	17921.1842	16422.4162	28046.5891
		others	1095.0789	1171.1275	3404.4495
PSNR	foreground area	ours	Inf	Inf	Inf
		others	24.7859	21.1464	36.8913
	background area	ours	5.5971	5.9764	3.6520
		others	17.7363	17.4448	12.8103
SSIM	foreground area	ours	1	1	1
		others	0.9662	0.9263	0.9969
	background area	ours	0.3989	0.6570	0.3147
		others	0.8996	0.9679	0.7238
E	foreground area	original image	7.3730	6.7845	5.4956
		ours	7.3730	6.7845	5.4956
		others	7.1222	6.2085	5.4711
	background area	original image	7.1647	6.2517	7.5699
		ours	2.8667	0.9031	3.2474
		others	6.596	5.7183	6.8473

The proposed algorithm achieved higher values in PSNR and SSIM, indicating lower signal distortion and higher structural similarity with the original images. Conversely, the background regions showed significant distortion, demonstrating that the method effectively preserves the details and structural information of the foreground while blurring the background. In contrast, the domain-adaptation fog synthesis algorithm resulted in significant signal distortion and structural differences between the foreground regions of the fogged images and the original images, failing to maintain the details and structural information effectively. Additionally, the background regions did not exhibit substantial distortion, indicating less effective privacy protection.

The results of the information entropy calculations show that, in the foreground region, the proposed algorithm does not cause a decrease in the image's information entropy value, while the domain-adaptation fog synthesis algorithm results in a slight decrease in entropy. In the background region, the proposed method significantly reduces the amount of information contained in the image, whereas the domain-adaptation fog synthesis algorithm does not achieve a substantial reduction in the image's information entropy.

6. Conclusion

This paper presents a novel image spatial fogging algorithm based on the ZoeDepth depth estimation

framework and the atmospheric scattering model. Firstly, the ZoeDepth framework is used for monocular depth estimation. Then, based on the atmospheric mathematical model, the fundamental theoretical model of the fogging algorithm is derived. The transmittance is calculated using the depth map and scattering coefficient, and the atmospheric light value is obtained using the dark channel prior algorithm, ultimately resulting in the fogged image. A comparative analysis is conducted with other fogging algorithms using four metrics: MSE, PSNR, SSIM, and Information entropy. The analysis results show that the foreground region has an MSE value of 0, a high PSNR value, an SSIM value of 1, and an unchanged information entropy ratio, while the background region shows opposite characteristics. This indicates that the proposed image spatial fogging algorithm provides high visual quality and realism in the foreground region while effectively blurring the background scene information to protect background privacy. This further validates the effectiveness and practicality of the proposed method, offering a more reliable and efficient solution for image fogging tasks in background privacy protection applications. In the future, this method can be applied in fields such as security surveillance, social media, and military security.

Acknowledgement

This work is supported in part by a grant from Education Technology Innovation project of Gansu Province (2021CYZC-22) and the Gansu Provincial Higher Education Institutions Industry Support Program (2023CYZC-19).

References

[1] Jizhe Zhou, Chi-Man Pun. Personal Privacy Protection via Irrelevant Faces Tracking and Pixelation in Video Live Streaming. IEEE Transactions on Information Forensics and Security 2021;16:1088-1103.

[2] Ruaa Sadoon Salman, Farah khiled Al-Jibory, Mauj Haider AbdAlkreem.Detect People's Faces and Protect Them by Providing High Privacy Based on Deep Learning. Tehnički glasnik2024;2024;18(1):92-99.

[3] Yifang Li, Nishant Vishwamitra, Hongxin Hu, Bart Knijnenburg, Kelly Caine. Effectiveness and Users' Experience of Face Blurring as a Privacy Protection for Sharing Photos via Online Social Networks. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2017;61(1):803-807.

[4] Mohammad Khojaste, Nastaran Moradzadeh, Ahmad Nickabadi. GMFIM: A generative mask-guided facial image manipulation model for privacy preservation. Computers & Graphics 2023;112(1):81-91.

[5] Sara Newman, Kannappan Palaniappan, Jianping Fan, Dan Lin. "Do You Know You Are Tracked by Photos That You Didn't Take": Large-Scale Location-Aware Multi-Party Image Privacy Protection. IEEE Transactions on Dependable and Secure Computing 2023;20(1):301-312.

[6] Jianping He, Bin Liu, Deguang Kong, Xuan Bao;, Na Wang, Hongxia Jin, George Kesidis. PUPPIES: Transformation-Supported Personalized Privacy Preserving Partial Image Sharing. IEEE/IFIP International Conference on Dependable Systems and Networks 2016;359-370.

[7] Nishant Vishwamitra, Yifang Li, Hongxin Hu, Kelly Caine, Long Cheng, Ziming Zhao, Gail-Joon Ahn. Towards Automated Content-based Photo Privacy Control in User-Centered Social Networks. Association for Computing Machinery 2022;65-76.

[8] Rakibul Hasan. Reducing Privacy Risks in the Context of Sharing Photos Online. Association for Computing Machinery 2020;1-11.

[9] Chau Yi Li, Ali Shahin Shamsabadi, Ricardo Sanchez-Matilla, Riccardo Mazzon, Andrea Cavallaro. Scene Privacy Protection. IEEE International Conference on Acoustics, Speech and Signal Processing 2019;2502-2506.

[10] Muhammad Bilal Sakha. Image Enhancement and Adversarial Attack Pipeline for Scene Privacy Protection. MediaEval Benchmarking Initiative for Multimedia Evaluation 2019.

[11] Chen Li, Weiqi Yan, Hongwei Zhao, Shihua Zhou, Yueping Wang. TFFD-Net: an effective two-stage mixed feature fusion and detail recovery dehazing network. Visual Computer 2024. https://doi.org/10.1007/s00371-024-03642-6

[12] Javed Aymat Husen Shaikh, Shailendrakumar Mahadev Mukane, Santosh Nagnath Randive. Lightweight progressive recurrent network for video de-hazing in adverse weather conditions. Visual Computer 2024. https://doi.org/10.1007/s00371-024-03683-x

[13] Bin Sheng, Ping Li, Yuxi Jin, Ping Tan, Tong-Yee Lee.Intrinsic image decomposition with step and drift shading separation. IEEE Transactions on Visualization and Computer Graphics 2018;26(2):1332-1346.

[14] Bin Sheng, Ping Li, Xiaoxin Fang, Ping Tan, Enhua Wu. Depth-aware motion deblurring using loopy belief propagation. IEEE Transactions on Circuits and Systems for Video Technology 2019;30(4):955-969.

[15] Yu Zhou, Zhihua Chen, Ping Li, Haitao Song, C. L. Philip Chen, Bin Sheng. FSAD-Net: feedback spatial attention dehazing network. IEEE transactions on neural networks and learning systems 2022;34(10):7719-7733.

[16] Codruta O. Ancuti; Cosmin Ancuti; Radu Timofte, "NH-HAZE: An Image Dehazing Benchmark with Non-Homogeneous Hazy and Haze-Free Images. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2020;1798-1805.

[17] Codruta O.Ancuti, Cosmin Ancuti, Mateu Sbert, Radu Timofte. Dense-Haze: A Benchmark for Image Dehazing with Dense-Haze and Haze-Free Images. IEEE International Conference on Image Processing 2019;1014-1018.

[18] [18]Cosmin Ancuti, Codruta O.Ancuti, Christophe De Vleeschouwer. D-HAZY: A dataset to evaluate quantitatively dehazing algorithms. IEEE International Conference on Image Processing 2016;2226-2230.

[19] Yanfu Zhang, Li Ding, Gaurav Sharma. HazeRD: An outdoor scene dataset and benchmark for single image dehazing. IEEE International Conference on Image Processing 2017;3205-3209.

[20] Codruta Orniana Ancuti, Cosmin Ancuti, Radu Timofte, Christophe De Vleeschouwer. I-HAZE: A Dehazing Benchmark with Real Hazy and Haze-Free Indoor Images. Advanced Concepts for Intelligent Vision Systems 2018;1804.05091.

[21] Md Nasim Khan, Mohamed M. Ahmed. Machine and Deep Learning Techniques for Daytime Fog Detection in Real Time with In-Vehicle Vision Systems Using the SHRP 2 Naturalistic Driving Study Data. Transportation Research Record 2023;2677(1):995-1011.

[22] Md. Imtiyaz Anwar, Arun Khosla. Classification of foggy images for vision enhancement. International Conference on Signal Processing and Communication 2015;233-237.

[23] Chi Zhang, Zihang Lin, Liheng Xu, Zongliang Li, Wei Tang, Yuehu Liu, Gaofeng Meng, Le Wang, Li Li. Density-Aware Haze Image Synthesis by Self-Supervised Content-Style

Disentanglement. IEEE Transactions on Circuits and Systems for Video Technology 2022;32(7):4552-4572.

[24] Christos Sakaridis, Dengxin Dai, Luc Van Gool. Semantic Foggy Scene Understanding with Synthetic Data.International Journal of Computer Vision 2018;126(9):973-992.

[25] Yading Zheng, Aizhong Mi, Yingxu Qiao, Yijiang Wang. Realistic Nighttime Haze Image Generation with Glow Effect. Association for Computing Machinery 2022;96–101.

[26] M.Colomb, K.Hirech, Philippe André, J.Boreux, P.Lacote, J.Dufour. An innovative artificial fog production device improved in the European project "FOG". Atmospheric Research 2008;87:242-251.

[27] Chuxuan Li, Ran Yi, Ali S G,et al. RADepthNet: reflectance-aware monocular depth estimation.. Virtual Reality & Intelligent Hardware 2022;4(5):418-431.

[28] Zizhuo Wang, et al. Robust blind image watermarking based on interest points. Virtual Reality & Intelligent Hardware 2024;6(4): 308-322.

[29] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, Rob Fergus. Indoor Segmentation and Support Inference from RGBD Images. European Conference on Computer Vision 2012;7576:746-760.

[30] Fayao Liu, Chunhua Shen, Guosheng Lin, Ian Reid. Learning Depth from Single Monocular Images Using Deep Convolutional Neural Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence 2016;38(10):2024-39.

[31] Rostam Affendi Hamzah, Haidi Ibrahim, Literature Survey on Stereo Vision Disparity Map Algorithms. Journal of Sensors 2016;2016(2):1-23.

[32] Wei Zhuo, Mathieu Salzmann, Xuming He, Miaomiao Liu. Indoor Scene Parsing with Instance Segmentation, Semantic Labeling and Support Relationship Inference. IEEE Conference on Computer Vision and Pattern Recognition 2017;6269-6277.

[33] Liang Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, Alan L.Yuille. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence 2018;40(4)834-848.

[34] S.Bhat, R.Birkl, Diana Wofk, Peter Wonka, Matthias Muller. ZoeDepth: Zero-shot Transfer by Combining Relative and Metric Depth. Advanced Concepts for Intelligent Vision Systems 2023;2302.12288.

[35] Kaiming He, Jian Sun, Xiaoou Tang. Single image haze removal using dark channel prior. IEEE Conference on Computer Vision and Pattern Recognition 2009;1956-1963.

[36] E.J.McCartney, Freeman F.Hall. Optics of the Atmosphere: Scattering by Molecules and Particles. Physics Today 1977;30(5):76–77.

[37] S.K.Nayar, S.G.Narasimhan. Vision in bad weather. Proceedings of the Seventh IEEE International Conference on Computer Vision 1999;2:820-827.

[38] Srinivasa G.Narasimhan, Shree K.Nayar. Vision and the Atmosphere. International Journal of Computer Vision 2002;48:233–254.

[39] Fan Guo, Jin Tang, Xiaoming Xiao. Foggy Scene Rendering Based on Transmission Map Estimation. International Journal of Computer Games Technology 2014;308629.

[40] Haoying Sun, Yutong Zheng, Qing Lang. Domain Adaptation for Synthesis of Hazy Images. Journal of Computer and Communications 2021;9:142-151.

[41] Zhou Wang, A.C.Bovik, H.R.Sheikh, E.P.Simoncelli. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13(4):600-612.