Single Underwater Image Enhancement Based on L_P-Norm Decomposition

JIANHUA WANG [1], HUIBIN WANG[1], GUOWEI GAO[2], HUIMIN LU [3], (Member, IEEE), AND ZHEN ZHANG[1]

[1]College of Computer and Information, Hohai University, Nanjing 210098, China
[2]School of Software Engineering, Anyang Normal University, Anyang 455000, China
[3]Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Kitakyushu 8048550, Japan

Corresponding author: Huibin Wang (hbwang@hhu.edu.cn)

ABSTRACT Integration of ocean monitoring networks with artificial intelligence has become a popular topic for researchers. Artificial intelligence plays an important role in underwater image processing. For optical images captured in an underwater environment, the light scattering and absorption effect caused by the water medium results in poor visibility, such as blur and color casts. A novel approach is proposed herein to enhance the single underwater image with poor visibility. Similar to other image enhancement strategies built on fusion principles, our method also generates two input channels from the original degraded image, and these two channels are modulated by their corresponding weight measures. However, the main innovation of our method is that we propose a new multilevel decomposition approach based on l_p-norm ($p = 0, 1, 2$) decomposition. According to the different sparse representation abilities of l_p-norm to an image's spatial information, our approach decomposes the image into three levels: detail level, structure level, and illuminance level. Thus, these three levels can be manipulated separately. Because this new decomposition approach is based on image structural contents, rather than direct per-pixel downsampling that is utilized in traditional multi-resolution pyramid decomposition, it is more accurate and flexible. Additionally, according to specific underwater imaging conditions, we carefully select two input channels and their three associated global contrast, local contrast, and saliency weight measures. Our method generates output with more accurate details and a better illuminant dynamic range. Generally, we are the first to impose an l_p-norm-based decomposition strategy on underwater image restoration and enhancement. Extensive qualitative and quantitative evaluations demonstrate that our strategy yields better results than state-of-the-art algorithms.

INDEX TERMS Image enhancement, l_p-norm, multilevel decomposition, underwater scene.
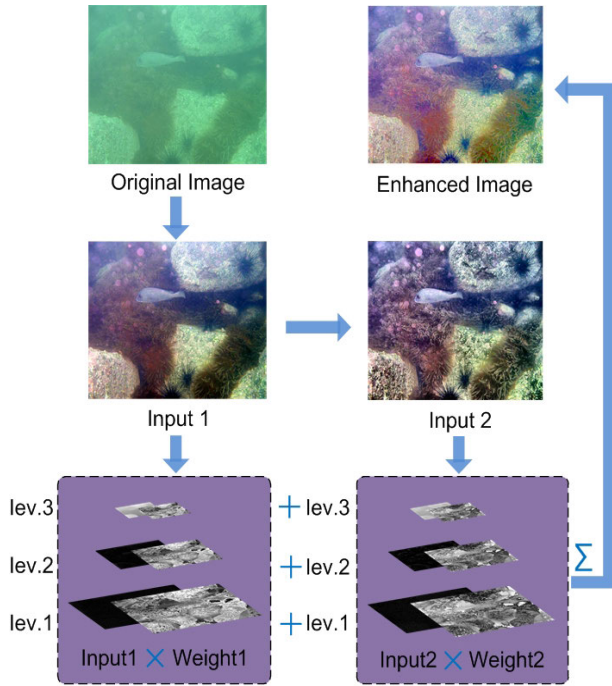
I. INTRODUCTION

Optical imaging technology is important in marine resource exploration and utilization (e.g., ocean scientific research, underwater environmental monitoring, underwater engineering construction and maintenance) [1]. In particular, owing to the integration of ocean networks with artificial intelligence, and the progression of artificial illumination technologies and high-quality sensor technologies [2], [3], marine observation techniques including underwater robots [4] are widely used, and research on underwater image processing has attracted more attention than before.

In underwater vision tasks, floating particles in the water attenuate light [4] owing to the scattering and absorption characteristics of the underwater environment. The scattering effect causes changes in the direction of light propagation, while the absorption effect reduces the energy [5]. Therefore, unlike images captured in the air, underwater images suffer from poor visibility: the fraction of light scattered from the water medium weakens the image contrast, and the absorption effect primarily causes chromatic casting. In such cases, typically used image features such as colour, intensity, and contours are not well characterized. In practice, when the

The associate editor coordinating the review of this manuscript and approving it for publication was Ke Gu [ID].

FIGURE 1. Primary framework of our l_p-norm-based decomposition algorithm. The original image is the picture that needs to be enhanced. Input1 is the histogram normalization colour-balanced version, and Input2 is derived through classical contrast local adaptive histogram equalization on Input1. Weight1 and Weight2 are Gaussian-normalized weight maps derived from weight measures of global contrast, local contrast, and saliency. Then, in both channels, the weight and input are multiplied within each level, and the results of the two channels are added together at corresponding levels. Finally, pyramid reconstruction is deployed by stacking all three levels after interpolation at lev.2 and lev.3.

sight distance is more than 10 m in common seawater, objects are almost unperceivable, and colors are substantially faded as light wavelengths are cut selectively [6].

To obtain better visibility in underwater images, we propose a novel single underwater image enhancement strategy based on a two-channel multilevel l_p-norm($p = 0, 1, 2$) decomposition. According to the different sparse expressive characteristics of l_0, l_1 and l_2 norm terms to the image's spatial information, we decompose the two input channels of a degraded underwater image into three levels: detail level, structure level, and illuminance level. We are the first to introduce this strategy into the restoration of underwater images. This approach outperforms the state-of-the-art algorithms, especially in terms of detail and illuminant dynamic range restorations. The primary framework of our l_p-norm-based decomposition algorithm is shown in Fig.(1), and the pseudocode is shown in Table(2).

The remainder of this paper is arranged as follows. In the rest of section I, the base knowledge of underwater light propagation and techniques that are related to underwater image enhancement are reviewed in detail. Then, our research work and innovation are briefly introduced. In Section II, we introduce our new single image l_p-norm decomposition-based dehazing technique in detail. Section III describes the experiments and analysis to illustrate the dehazing effects of

our strategy. Section IV briefly summarizes and concludes our approach. Finally, our acknowledgements are presented in section V.

A. UNDERWATER LIGHT PROPAGATION

For the underwater scene, light strength, colour and direction are significantly affected by the water medium due to absorption and scattering [6]. Because the absorption effect varies with light wavelength, water-induced shortwavelength light (green and blue) casts more than longwavelength light (red). In practice, the attenuation and colour loss also depend on the total distance between the observer and the scene. Meanwhile, the scattering effect results in poor contrast of underwater scenes and render the image obscure and misty.

Based on the research of McGlamery [7] and Jaffe [8], in such underwater degraded scenes, only part of the reflected light can reach the camera lens, and the total irradiance incident on the pixel point of an image plane in a camera contains three primary components: a direct component, forward scattering, and back-scattering. Practically, the back-scattering effect is the principal source of image structure (contrast) loss and colour casting in underwater scenes.

At each pixel point x of an underwater image, the direct component is

$$I_d(x) = I_{obj}(x)e^{-\eta d(x)}, \tag{1}$$

where $I_{obj}(x)$ is the intensity of light radiance from the target object, $d(x)$ is the distance from the object to the camera, and η is the attenuation coefficient of the underwater medium. The exponential term $e^{-\eta d(x)}$ is known as the transmission of the underwater medium and depicts the amount of light transmitting from the target object to the camera.

The back-scattering component is

$$I_{bs}(x) = V_{bs}(x)(1 - e^{-\eta d(x)}), \tag{2}$$

where $V_{bs}(x)$ is the colour vector known as back-scattering light.

For each pixel point x, the total irradiance incident light intensity $I(x)$ is primarily described by these two additive components:

$$\begin{aligned} I(x) &= I_d(x) + I_{bs}(x) \\ &= I_{obj}(x)e^{-\eta d(x)} + V_{bs}(x)(1 - e^{-\eta d(x)}) \end{aligned} \tag{3}$$

B. RELATED WORK

Underwater image enhancement is a fundamental process for underwater vision tasks. The existing restoration and enhancement techniques can be categorized into hardware-based and software-based methods:

First, an important class of enhancement approaches uses specialized hardware [1] (e.g., light field cameras, polarization-based methods, and lidar imaging (UWLI) systems). In [9], D. G. Dansereau et al. demonstrated the potential use of light field cameras in underwater dehazing. Based on the light field camera, Skinner and Johnson-Roberson [10]

proposed a physical model in which image structural information was embedded. It recovered the dehazed image and estimated the range-dependent transmission. In [11], a polarization filtering approach that restored visibility while adaptively suppressing noise was proposed. Partial polarization-based techniques use several images of different degrees of polarization; they are effective in recovering distant regions but not applicable to dynamic scenes. In [12], the divergent-beam UWLI system was used to capture turbid underwater images. Generally, these hardware acquisition systems are complex, expensive, power consuming [6], and difficult to operate, especially in an underwater environment.

Software-based strategies can be separated into single- or multi-image methods. A single input image strategy contains light propagation model-based, contrast-based, statistical-based, dark channel prior (DCP), and super-resolving (SR) methods. The light model approach models light as the product of the illumination and reflectance elements. The illumination element typically contains low-frequency components. The reflectance element includes high-frequency information, such as edge and texture. Kimmel *et al.* [13] proposed a variational low-pass filter to estimate the illumination elements and achieved visually pleasing results. Ancuti *et al.* [14] reconstructed the intrinsic brightness by Koschmieder's visibility model [15] in heterogeneous lighting conditions. For the contrast method, Tan [16] proposed a dehazing algorithm based on observations: enhanced images typically exhibit higher contrast and light changes smoothly in a small local area. Their algorithm can enhance visibility, but it assumes that the depth map must be smooth except along edges [17]. Gu et al. developed the image quality assessment (IQA) model and contrast changed image database (CCID2014) and then presented the contrast enhancement strategy using the histogram modification framework [18], [19]. Gu *et al.* [20] presented an automatic contrast enhancement strategy based on the concept of saliency preservation. For statistical analysis, Fattal [21] employed a statistical graphical model and solved the ambiguity of light colour. Because scene albedo and depth are two statistically independent components, Nishino *et al.* [22] proposed a factorial Markov random field Bayesian defogging algorithm. In addition, Liu et al. utilized maximum posterior with prior propagation models and residual convolutional neural network (RCNN) techniques to predict propagation direction, which is more robust to unwanted local minimums [23]. However, there is a lack of specialized underwater datasets, so the performance for underwater scenes is limited.

Several image dehazing algorithms are based on DCP [24]. Originally, the DCP method was applied for image dehazing for airlight scenes. It defines the regions of small transmission as those with large minimal values of colours [6]. Gibson et al. demonstrated mathematically that the DCP theory performed well for image defogging [26]. Chen *et al.* [27] proposed a d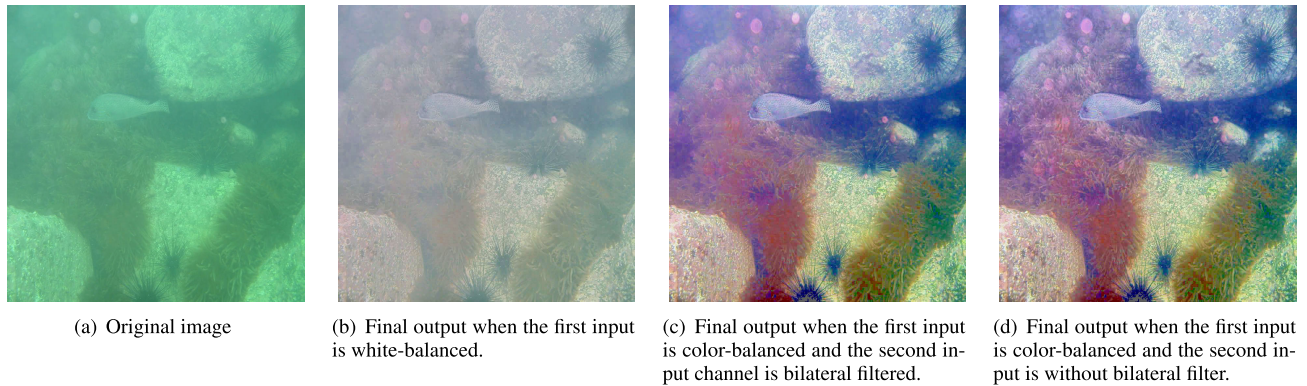ehazing method to suppress artefacts by minimizing the gradient residual. Pei and Lei [28] and Chen *et al.* [29] used the DCP theory to restore a night-time haze image and near-infrared image, respectively. Huang *et al.* analyzed the original DCP defogging algorithm for adapting to various weather conditions and proposed an improved DCP algorithm with depth estimation, colour analysis, and visibility restoration modules [30]. In underwater imaging conditions, Chiang and Chen [31] separated the foreground and background regions based on DCP and removed the haze and colour distortion based on this segmentation. Paulo Drews *et al.* [32] assumed that the primary source of underwater visual information was from blue and green colour channels, and they proposed an underwater dark channel prior (UDCP) method that estimated the transmission map better than the conventional DCP. In general, the DCP strategy can generate relatively pleasing visibility, but it is usually related to an erosion problem [17].

Another category of underwater image dehazing algorithms is super-resolution (SR) descattering. Many studies have shown that image structures and details repeat themselves [33]. Based on this observation, many HR images can be restored from self-examples [34]–[36]. Zhang *et al.* [37] designed an end-to-end deep neural texture transfer SR model to extract more texture details from Ref images, but datasets of underwater scenes are insufficient. Most image SR algorithms are based on the 'smoothness' assumption that is similar to the classic optical flow theory. These methods result in artefacts and blurring at image edges and corners.

In general, single underwater image enhancing techniques are commonly straightforward to implement. However, compared with the more robust multi-resolution fusion strategy, they often suffer from problems such as image structure loss, ghosting halo artefacts, and spurious edges.

The method of multiple image restoration utilizes several images as input. These input images are captured at the same scene but in different conditions, such as different lighting [38], [39], and different degrees of polarization [11]. In addition, some supplemental information about scene models can be referenced. Kopf *et al.* [40] used the existing digital terrain and urban models to enhance images. These different medium properties and conditions may provide different important hazy image information and ultimately output pleasing visibility. However, in typical cases, the acquisition operation for multi-view images is time consuming and difficult to perform. Furthermore, because these images and scene model information are unavailable, these methods are impractical for general users.

Among these different image enhancement and restoration approaches, multi-resolution pyramid decomposition and fusion is a typical strategy. It is based on direct downsampling at a pixel scale [6]. This method can weaken the halo and artefact effects and yield relatively good restoration results. However, it causes the loss of spatial structure information, especially textures and details, which are of key importance to the enhancement of underwater images.

| (a) Original image | (b) Final output when the first input is white-balanced. | (c) Final output when the first input is color-balanced and the second input channel is bilateral filtered. | (d) Final output when the first input is color-balanced and the second input is without bilateral filter. |

FIGURE 2. Comparison of different processes of the two input channels. (a) The original degraded image needs enhancement. (b) is the final output when the first input channel is white balanced; apparently, it is too white. (c) is the final output when the first input is colour-balanced, and the second input channel is bilateral filtered; compared to (b), colour visibility is improved conspicuously. (d) is the final output derived from the colour-balanced version of the first input and version of the second input that removed the bilateral filter. (d) exhibits almost the same colour appearance as (c) but with more details.

C. OUR METHOD AND INNOVATION

We propose a novel two-channel multilevel decomposition approach based on the l_p-norm ($p = 0, 1, 2$). According to different sparse expressive traits of the l_0, l_1 and l_2 norm terms to spatial information such as edges and structures, we decompose both input channels derived from one degraded underwater image into three levels: detail level, structure level, and illuminance level. The downsampling process was performed only on the structure level and illuminance level, which are not sensitive to spatial resolution (see Fig.(5)(b)(c) and (e)(f)).

We are the first to introduce this l_p-norm-based decomposition strategy into underwater image enhancement. The L_p-norm decomposition method was used for tone mapping to compress the dynamic brightness range in [41]. Interestingly, we discover that this strategy not only improves illuminance appearance can also extract detail and structure information. Therefore, we introduce this l_p-norm to our research work. However, its simple and direct implementation is insufficient, so we generate the first histogram normalization colour-balanced input and the second contrast local adaptive histogram equalized input from a degraded single underwater picture as the two input channels. Additionally, we carefully select features such as global contrast, local contrast, and saliency as the weight measures to modulate these two input channels. We removed the chromatic weight but placed more emphasis on the contrast information. Our l_p-norm-based decomposition strategy includes two procedures: rough decomposition and refined decomposition. Therefore, it is more accurate than traditional pyramid decomposition. Furthermore, our decomposition method is only performed at the grey channel, and it is simpler than the traditional pyramid approach that executes at all the R, G, and B channels. The detailed contents are presented in the following sections.

II. L_P-NORM(p = 0,1,2) DECOMPOSITION-BASED UNDERWATER ENHANCEMENT

The fundamental idea of the decomposition algorithm is to combine several inputs that are tailored by corresponding weights. To obtain good visibility of the final output image, the first procedure for designing a decomposition algorithm is to select appropriate inputs and weights. Inputs and weight maps are chosen to maintain the most significant features of the original image, and they are application dependent.

The second key step is to decompose and merge these modulated inputs. We introduce a novel multilevel decomposition strategy that does not resort to deriving the inputs based on the scene physical model and naive multi-resolution pyramid model. Instead, we aim for a l_p-norm-based multilevel decomposition technique in accordance with the different sparse expressive capacities of the l_0, l_1 and l_2 norm terms to image spatial information. The detailed discussion and analysis are depicted as follows:

A. INPUTS AND WEIGHT MAPS

1) THE TWO INPUT CHANNELS

Our approach exploits two input channels (please refer to Fig.(1)) that are derived from the single degraded underwater image. For **the first input**, we resort to the colour correction technique to depress the effect of colour casting in a water medium. The colour correction technique aims to enhance image visibility by discarding unwanted colour casts caused by light attenuation and various illuminant conditions in underwater scenes. We choose a histogram normalization colour balance strategy that is based on histogram normalization. Compared to our work, the authors of [6] performed a white balancing correction. However, the result of white balancing correction tends to appear too white in general. Thus, gamma correction was implemented to increase the difference between darker and lighter regions. Although this strategy is effective, it is slightly more complex than our methods. Furthermore, our method can enhance chromatic visibility while avoiding the gamma correction's drawback of losing contrasts in under-/over-exposed regions. Fig.(2)(b) is the final output derived from the white balanced version of the first input channel. It is obviously too white. (c) demonstrates that the simple colour balancing approach yields notably

better colour visibility than the white balanced approach (b) and decreases the gamma correction disadvantage of losing contrasts.
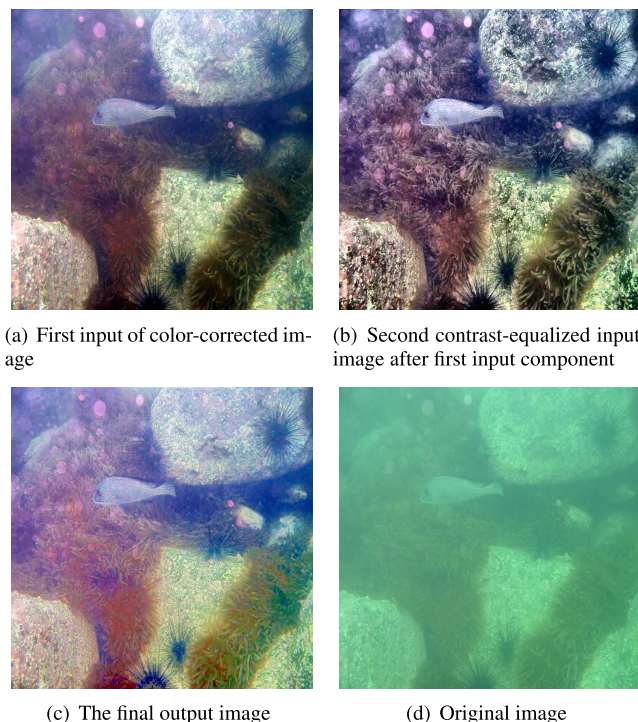
In the first input channel, we discarded the colour shift. However, in the turbid water medium, underwater scenes still present significant noise and lack of contrast owing to light absorption and scattering. Therefore, **the second input** is designed from this first input to enhance the contrast.

In [5], the author chose the temporal bilateral filter to denoise and smooth the underwater image. However, we found that this bilateral filter indeed reduced the noise level, but at the cost of losing many small details. This weakens the defuzzification effect of our enhancement approach. As shown in Fig.(2)(c) and (d), the final output image generated from the colour-balanced first input and second input with no bilateral filter presents more detailed information than (c). Additionally, with an acceptable noise level owing to the denoising benefit of our l_p-norm decomposition strategy, we decided to remove the bilateral filter.

Subsequently, to achieve an optimal contrast level after the first input component, the second input is designed through classic contrast local adaptive histogram equalization. Generally, the typical global operators can be applied [5], but they must be either specified by the user or estimated from an input image. Furthermore, the improvements acquired by these operators are at the expense of the remaining regions. Therefore, we choose the local adaptive histogram. This technique functions automatically with minor distortions. The contrast between adjacent structures is portrayed maximally such that it occupies a larger portion of the intensity range than the original image. As shown in Fig.(3), after the colour correction, chromatic appearance of the first input channel (a) achieved significant improvement compared with the original image (d). (b) is the contrast-enhanced image using the local adaptive histogram contrast equalization method after the chromatic operator of the first input channel (a). It exhibits a similar colour as (a) but with a signal-highlighted contrast. Therefore, (b) is defined as the second input channel of our decomposition strategy.

2) WEIGHT MAPS

Because of the absorption and scattering phenomena of light propagating in turbid water, underwater images often appear obscure and colour cast. Therefore, weight measures such as chromatic, contrast, and luminance are strongly correlated with image restoration in underwater scenes [6]. For every corresponding pixel point in those two input channels, a relatively higher weight value appears in the final image with a larger proportion. We utilized simple colour balance correction to adjust the chromatic features on both input channels, so it is unnecessary to choose the chromatic weight. Therefore, in comparison with a previous study [6], we emphasize contrast and luminance weights such as local contrast, global contrast, and salience features.

(a) First input of color-corrected image

(b) Second contrast-equalized input image after first input component

(c) The final output image

(d) Original image

FIGURE 3. Demonstration for the two input channels. (a) is the first input channel with a histogram normalization colour-balanced approach. Compared with the original image (d), the colour appearance of (a) is improved. (b) is the second input channel using the classic contrast local adaptive histogram equalization method (a). Hence, (b) exhibits similar colour appearance as (a), but with highlighted contrast. (c) is the finally dehazed output image.

a: GLOBAL CONTRAST WEIGHT

This weight measures the global contrast of the input. It computes the absolute value of the Laplacian filter on each input luminance channel. The Laplacian filter is a high-pass filter that can extract edges and textures and assign high values to them. However, this global Laplacian contrast weight cannot effectively distinguish between a ramp and a flat region. Thus, using only this weight is insufficient for recovering the contrast of dim and obscure underwater scenes. To fully describe the structure information, we select the local contrast weight as another weight factor.

b: LOCAL CONTRAST WEIGHT

The local contrast weight demonstrates each pixel and its average neighborhood's variation that is primarily sensitive to the transitions between the highlighted and shadowed regions. Therefore, this weight measure can enhance the local contrast of the underwater image. It is computed as the standard deviation between the pixel luminance level and the mean value of its local surrounding region:

$$W_{LC} = \|I^k - I^k_{\omega hc}\|, \qquad (4)$$

where W_{LC} represents the local contrast weight, I^k represents the luminance channel of the input image, and $I^k_{\omega hc}$ denotes the low-pass version of I^k [5].

c: SALIENCY WEIGHT

Based on image features such as intensity, colour, or orientation, the saliency weight map estimates the contrast of the image regions relative to their surroundings and assesses whether a certain object is more prominent than the remaining images. To improve visibility of the objects that lose their prominence in underwater scenes, we exploit this saliency weight. We opted for the biological centre-surround inspired and computationally efficient approach of Achanta *et al.* [42]. In general, this spatial structure and illuminance sensitive saliency map enhances the large regions, estimates uniform values for the whole salient regions, and increases the contrast in highlighted and shadowed regions.

For the two input channels discussed previously, we let the normalized weight values $\bar{W}^k = W^k/(\sum_{k=1}^{K} W^k)$ by constraining $\sum_{k=1}^{K} \bar{W}^k = 1$, where $K = 2$ denotes the two input channels. W^1 and W^2 represent the sum of the global contrast weight, local contrast weight, and saliency weight derived from both input channels.

B. L_P-NORM(p = 0,1,2) BASED DECOMPOSITION

1) NAIVE MULTI-SCALE METHOD

The single-scale fusion approach directly fuses the inputs with the corresponding normalized weight maps to conserve the most significant features. However, this method creates undesirable halo artefacts primarily in locations characterized by strong transitions of the weight maps. Subsequently, multiscale decomposition, such as [5], [43]–[45], is selected. Among them, the Laplacian pyramid decomposition is the typically used method:

$$R_l(x, y) = \sum_{k=1}^{K} \left\{ G_l \left\{ \bar{W}^k(x, y) \right\} L_l \left\{ I^k(x, y) \right\} \right\}, \quad (5)$$

where l represents the different levels of the fusion, typically $l = 5$. $R(x, y)$ is the enhanced image version, (x, y) is the pixel point coordinate, and k is the index of the two inputs: $I^k(x, y), k = 1, 2$. $\bar{W}^k(x, y)$ is the normalized weight maps, $G_l\{\bar{W}^k(x, y)\}$ is the Gaussian filtering for weight $\bar{W}^k(x, y)$ at the l_{th} level, and $L_l\{I^k(x, y)\}$ is the Laplacian filtering output for $I^k(x, y)$.

For the Laplacian term $L_l\{I^k(x, y)\}$, the Laplacian pyramid filters the input image $I^k(x, y)$ using a low-pass Gaussian kernel at the l_{th} level and downsamples the filtered image by a factor of two. It subsequently subtracts from the original input $I^k(x, y)$ with an upsampled version of this low-pass image, such that the high-pass image components remain. Therefore, it approximates the (inverse of the) Laplacian and uses the decimated low-pass image as the input for the subsequent $(l + 1)_{th}$ level of the pyramid. The resulting representation, i.e., the Laplacian pyramid, is a set of quasi-bandpass versions of the image.

Similarly, for each normalized weight map $\bar{W}^k(x, y)$, a Gaussian pyramid is computed. Considering that both the Gaussian and Laplacian pyramids have the same number of layers, combining the Laplacian input images and Gaussian-normalized weights is independently performed at each level, and yields the fused pyramid depicted in Eq. (5).

This procedure is iterated successively for each l_{th} pyramid level in a bottom-up manner. The final merged image $R(x, y)$ is reconstructed by summing the resulting inputs from Eq. (5) at all l_{th} levels of the pyramid:

$$R(x, y) = \sum_{l} R_l(x, y) \uparrow^d, \quad (6)$$

where $\uparrow^d$ is the upsampling operator with a factor of $d = 2^{l-1}$.

This naive multiscale fusion is motivated by the human visual system. This system is highly sensitive to sharp transitions in smooth image patterns but less sensitive to variations/artefacts occurring on edges and textures [6]. Benefitting from this fusion strategy at every scale level, the potential halo artefacts owing to the sharp transitions can be minimized. However, this direct pixel scale downsampling strategy results in the loss of spatial resolution and thus, the loss of image information. Therefore, we introduce the l_p-norm-based decomposition approach. Indeed, the conception of the l_p-norm has been implemented for image tone mapping in [41]. However, it is interesting that the l_p-norm term has the extractive capability to structure and detail information of an image in addition to compressing the illuminance dynamic range. A detailed discussion is presented in the following sections.

2) SPARSE TRAITS OF L_P-NORM(p = 0,1,2)

A sparse regularization operator can remove some features with useless information by learning, i.e., automatic feature selection and dimension reduction. Based on sparse dimension reduction, the original model can be easier to understand. In psychology research [46], [47] as an example, it was found that human vision is more sensitive to structural information such as edges and textures. This visual mechanism facilitates capturing the primary semantic information of the scene. Therefore, in intrinsic decomposition [48], [49], it is typically assumed that the structural information in the reflectance layer (a concept similar to the detail layer) is sparse [41]. The l_p-norm ($p = 0, 1, 2$) is a method for achieving sparse regularization. The l_0-norm refers to the number of non-zero elements in the matrix. If a parameter matrix W is normalized with its l_0-norm, it is desirable for its elements to be zero. The l_1-norm refers to the sum of absolute values of each element in a matrix, which is also called the Lasso regularization. l_1-norm is the optimal convex approximation of the l_0-norm. The l_2-norm is the sum of the squares of each element in the matrix, followed by its square root. We apply the minimized rule term of l_2-norm $\|W\|_2$ such that every element of W is small and close to zero.

Compared with the l_2-norm, the l_1-norm tends to produce fewer features, and other features are zero. The l_2-norm selects more features that are close to but not equal to zero. Therefore, the l_1-norm is highly useful in feature selection. The l_2-norm is often used for regression (or fitting).

It can restrict the model space, thereby avoiding the under-fitting or over-fitting problem and improving the model's generalization ability. In addition, for optimization or the numerical calculation question, the l_2-norm can address the matrix processing difficulty when the condition number is not sufficient. The l_1-norm sparsity term exhibits the outlier-rejection nature. It preserves the edges with large gradients in an image, and its piecewise smoothness nature results in relatively weak structural information. In contrast, the l_0-norm term has piecewise flattening effects, and it can force small textural gradients of the detail layer to be zeros while leaving the primary detail gradients intact. It is noteworthy that the piecewise flattening property of the l_0-norm term has denoising effects to some extent. Therefore, the l_0-norm gradient sparsity term is used to model the detailed information, which is enhanced. Additionally, to reduce the halo artefacts, a l_1-norm gradient sparsity term is imposed to preserve the edges.
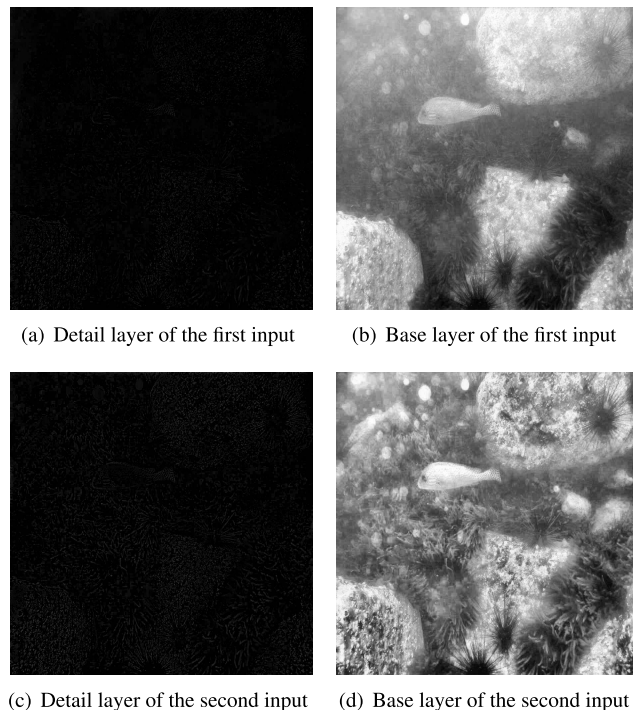
Consequently, the l_p-norm$(p = 0, 1, 2)$ based multilevel decomposition [41] is introduced to our research work. Our multilevel algorithm decomposes the image into three levels: detail level, structure level, and illuminance level. Hence, different attributes of a single image can be denoted by different layer levels with the l_p-norm and can, therefore, be manipulated differently to yield a more flexible and effective enhancement strategy. Generally, our l_p-norm-based decomposition framework involves two steps: rough-level and refined-level decompositions. The l_2-norm is applied to force the base layer to be close to the input image and avoid the over/under regression problem. For the rough-level decomposition, the l_0-norm term is imposed on the detail layer to model the detail prior, and an l_1-norm term is imposed on the base layer to preserve structures and reduce halo artefacts. For the rough-level decomposition, structure and illuminance information are relatively weaker spatial features than detailed information. Therefore, we impose the l_1-norm term to separate structure information with the local mean brightness, i.e., illuminance information.

3) ROUGH-LEVEL DECOMPOSITION

According to the traits of the l_p-norm(p = 0,1,2) analyzed in the last section, rough-level decomposition is performed based on the following model:

$$\hat{B}_1 = \arg\min_{B_1} \sum_k \left\{ \|I - B_1\|_2^2 + \theta_1\|\nabla B_1\|_1 + \theta_2\|\nabla D_1\|_0 \right\},$$

$$(7)$$

where I denotes the original input image, B_1 is the base layer and D_1 represents the detailed layer of the rough decomposition. k is the index of every pixel point of I, and ∇ is the derivative operation at pixel $k(x, y)$. This derivative operation utilizes gradient information to depict detailed information such as edges and textures. $\|I - B_1\|_2^2$ is a l_2-norm term that constrains the base layer B_1 to be close to the original image

(a) Detail layer of the first input

(b) Base layer of the first input

(c) Detail layer of the second input

(d) Base layer of the second input

FIGURE 4. Output of the first rough-level decomposition based on L_0, L_1, and L_2 norms. (a) and (c) are the corresponding detail layers of the two input channels, (b) and (d) are the corresponding base layers. Owing to space limitation, images (a) and (c) may not be sufficiently clear. Please refer to Fig.(5) (a) and (d) for clearer versions.

I. The term $\|\nabla B_1\|_1$ is the l_1-norm of layer ∇B_1, and $\|\nabla D_1\|_0$ represents the l_0-norm of layer ∇D_1.

The right side of Eq. (7) can be written as

$$\arg\min_{B_1} \sum_k \left\{ (I - B_1)^2 + \theta_1 \left| \frac{\partial B_1}{\partial(x, y)} \right| + \theta_2 I_{ind}\left(\frac{\partial D_1}{\partial(x, y)} \right) \right\},$$

$$(8)$$

where ∂ is the partial derivative operation of pixel point k along the directions of coordinate axes x and y. $I_{ind}(x)$ is the function that indicates whether the independent variable belongs to the group. In our case, if $x \neq 0$, the dependent variable is one; otherwise, it is zero.

In Fig.(4) that is generated from this first rough decomposition model, (a) and (c) are the l_0-norm related detail layers D_1. D_1 is calculated by $D_1 = I - B_1$ and contains the primary texture information. (b) and (d) are the base layers B_1 that correspond to the l_1-norm and contains information on the structure and illuminance. Additionally, (a) and (b) are generated from the colour-balanced first input channel in Fig.(3)(a). (c) and (d) are derived from the colour-balanced and contrast-equalized second input channel in Fig.(3)(b). Obviously, only this rough-level decomposition cannot effectively classify the structure and illuminance information. Therefore, we propose the refined decomposition strategy on base layer B_1 in the next section to more accurately extract different components of the underwater image.

4) REFINED-LEVEL DECOMPOSITION

In refined-level decomposition, we aim to separate structure and illuminance information from the base layer that is generated from the rough decomposition. The structure and illuminance levels are relatively weaker spatial scales than the detailed information. Additionally, considering the piecewise smoothness characteristic of the l_1-norm and the piecewise flattening characteristic of the l_0-norm, we decide to choose the l_1-norm as the model of refined-level decomposition:

$$\hat{B}_2 = \arg \min_{B_2} \sum_k \left\{ \|I - B_2\|_2^2 + \theta_3 \|\nabla B_2\|_1 \right\} \quad (9)$$

$$= \arg \min_{B_2} \sum_k \left\{ (I - B_2)^2 + \theta_3 \left| \frac{\partial B_2}{\partial(x,y)} \right| \right\}. \quad (10)$$

Fig.(5) demonstrates that after decomposition with Eq. (8) and Eq. (10), both the first colour-balanced input channel and the second contrast-equalized channel include three levels: the detail (or texture) level shown in Fig.(5)(a) and (d), the structure level shown in Fig.(5)(b) and (e), and the illuminance level shown in Fig.(5)(c) and (f). Furthermore, (a) and (d) contain the texture detail elements similar to Fig.(4)(a) and (c), while the structure and illuminance information disassembled from the single base layer of the rough-level decomposition in Fig.(4) are also depicted herein separately.

The derived detail level, structure level, and illuminance level can be denoted by D_t, S_t, and I_l, respectively. Subsequently, to adjust (including stretch or compress) these three components according to different underwater scenes, we add three coefficients: ξ_d, ξ_s, and ξ_i on three levels separately.

$$L_p \left\{ I^k(x,y) \right\} = \left\{ \xi_d D_t^k, \quad \xi_s S_t^k, \quad \xi_i I_l^k \right\}, \quad (11)$$

where k is the index of the two inputs: $I^k(x,y), k = 1, 2$. $L_p\{I^k(x,y)\}$ is the three level l_p-norm based decomposition for $I^k(x,y)$.

5) FUSION

After the l_p-norm-based decomposition of the two input channels, the fusion process is necessary for achieving the final clear underwater image. We adopted the following method as the fusion strategy:

$$R_l(x,y) = \sum_{k=1}^K \left\{ G_l \left\{ \bar{W}^k(x,y) \right\} L_p \left\{ I^k(x,y) \right\} \right\}, \quad (12)$$

where l represents the different fusion levels, here, $l = 3$. $R(x,y)$ is the enhanced output, (x,y) is the pixel point coordinate, and k is the index of the two inputs: $I^k(x,y), k = 1, 2$. $G_l\{\bar{W}^k(x,y)\}$ is the Gaussian filtering for the normalized weight $\bar{W}^k(x,y)$ at the l_{th} level, and $L_p\{I^k(x,y)\}$ is the three level l_p-norm based decomposition for $I^k(x,y)$.

In Fig.(5), (g), (h), and (i) are Gaussian-normalized weight maps of the colour-balanced first input at the three different levels, and (j), (k), and (l) are Gaussian-normalized weight

maps of the contrast-equalized second input at three levels. Three levels of the two inputs are both $1/2$ down sampled stepwise from the bottom (g) and (j) to the top (i) and (l). Additionally, the structure level (b) and (e) and illuminance level (c) and (f) are not as sensitive to the image spatial resolution as the detail level (a) and (d). Thus, $1/2$ down-sampling is adopted on the structure and illuminance levels. We magnified (a), (b), (c), (d), (e), and (f) deliberately to depict the small textures more clearly; in fact, (a), (d), (g), (j), (b), (e), (h), (k), (c), (f), (i), and (l) are of the same sizes.

6) MODEL SOLUTION

Model (7) can be solved by the alternating direction method of multipliers (ADMM). ADMM is an optimization method for constrained problems. By decomposing the coordination process and using the alternating solution, the large global problem is decomposed into several small and easy-to-solve local subproblems to obtain the solution to the large global problem by coordinating the subproblems.

The matrix-vector form of Eq. (7) is

$$\arg \min_b \sum_k \left\{ \|i - b\|_2^2 + \theta_1 \|\nabla b\|_1 + \theta_2 1^\top I_{ind}(\nabla d) \right\}, \quad (13)$$

where $i, b, and d, \in R^N$ are the concatenated vector forms of I, B, and D in Eq. (7), respectively. $1 \in R^{2N}$ is a vector of all ones. The resultant augmented Lagrangian function in our model is

$$\mathcal{L}(b, c_1, c_2, y_1, y_2)$$
$$= \frac{1}{2} \|i - b\|_2^2 + \theta_1 \|c_1\|_1 + \theta_2 1^\top I_{ind}(c_2) + (c_1 - \nabla b)^\top y_1$$
$$+ (c_2 - \nabla d)^\top y_2 + \frac{\rho}{2} (\|c_1 - \nabla b\|_2^2 + \|c_2 - \nabla d\|_2^2), \quad (14)$$

where $c_1, c_2 \in R^{2N}$ are introduced to replace ∇b, and ∇d respectively. y_i, and $i = 1, 2$ are the Lagrangian dual variables. At iteration step m, the function (14) is optimized by minimizing several primal subproblems and maximizing the dual problems alternatively:

1. For b^{m+1}:

We solve b^{m+1} through an FFT transformation by splitting vector $c_1, c_2, y_1, and y_2$ into $(*)_{(*,1)}^{m+1}$ and $(*)_{(*,2)}^{m+1}$, and transform the objective function with respect to b^{m+1} into a quadratic programming problem.

2. For c_1^{m+1}:

c_1^{m+1} can be solved through soft shrinkage:

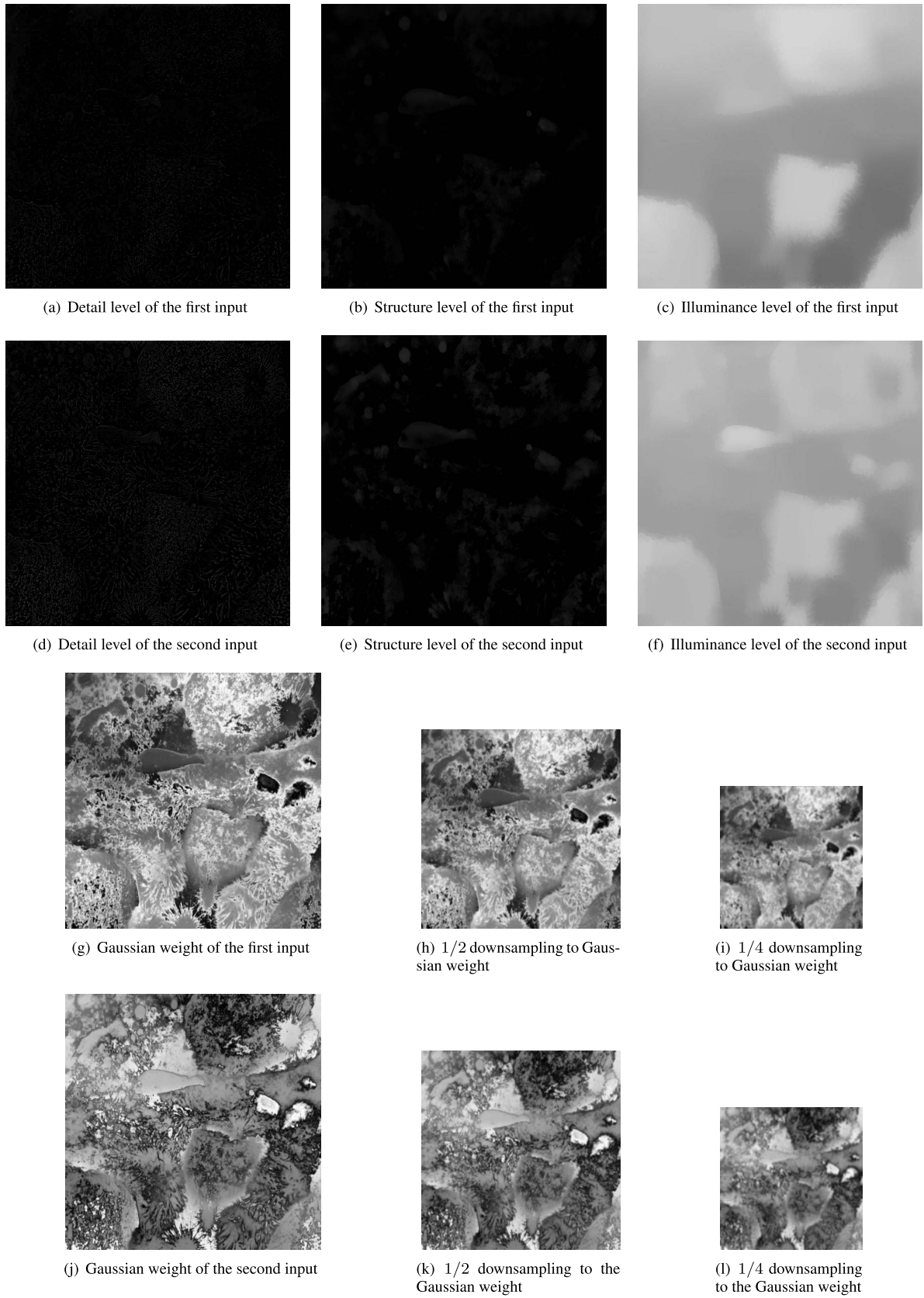$$c_1^{k+1} = \Gamma_{\lambda_1/\rho^m}(\nabla b^{m+1} - y_1^m/\rho^m), \quad (15)$$

where $\Gamma(x)$ is the soft thresholding function.

3. For c_2^{m+1}:

The objective function with respect to c_2^{m+1} can be solved in a per-entry manner. We denote by subscript j the j_{th} entry of a vector. The solution at entry j is

$$c^{m+1}{}_{2,j} = \begin{cases} 0, & \text{if } (f_j^m)^2 \leq \frac{\lambda_2}{\rho^m}; \\ f_j^m, & \text{otherwise.} \end{cases} \quad (16)$$

(a) Detail level of the first input

(b) Structure level of the first input

(c) Illuminance level of the first input

(d) Detail level of the second input

(e) Structure level of the second input

(f) Illuminance level of the second input

(g) Gaussian weight of the first input

(h) 1/2 downsampling to Gaussian weight

(i) 1/4 downsampling to Gaussian weight

(j) Gaussian weight of the second input

(k) 1/2 downsampling to the Gaussian weight

(l) 1/4 downsampling to the Gaussian weight

FIGURE 5. Different decomposition levels and the corresponding Gaussian-normalized weight maps.

| (a) Original image | (b) 1.0, 0.4, 0.7 | (c) 2.0, 0.4, 0.7 | (d) 1.0, 1.5, 0.7 | (e) 1.0, 0.4, 0.3 |

FIGURE 6. Comparison of different values of parameters (ξ_d, ξ_s, ξ_i) in Eq. (11). To demonstrate clearly, we use two images, which are shown in the top and bottom rows. ξ_d adjusts the detail level, ξ_s adjusts the structure level, and ξ_i affects the illuminance level. (a) are the two original degraded images that require enhancement. (b) shows the output pictures when (ξ_d, ξ_s, ξ_i) are set to (1.0, 0.4, 0.7). In (c), the detail level is amplified twice that of (a). Obviously, the textures and details suffer from overboosting effects. (d) shows the case of (1.0, 1.5, 0.7), in which the structure level is amplified nearly four times. Thus, the restored images are blurred, and many details are lost. In (e), the illuminance level parameter ξ_i is set to 0.3, and the derived outputs are dark and obscure.

where

$$f_j^m = (\nabla(i - b^{m+1}) - y_2{}^m/\rho^m), \quad j = 1, \ldots, 2N. \quad (17)$$

4. Dual ascent for Lagrangian multipliers.
5. Update ρ^{m+1} as $\rho^{m+1} = 2\rho^m$.
After b is obtained, d can be calculated by $d = i - b$.

For Eq. (9) and Eq. (10):

$$\hat{B} = \arg\min_B \sum_k \left\{ \|I - B\|_2^2 + \theta_1 \|\nabla B\|_1 \right\}$$

$$= \arg\min_B \sum_k \left\{ (I - B)^2 + \theta_1 \left| \frac{\partial B}{\partial(x, y)} \right| \right\},$$

where we use $\theta_2 = 0$ in Eq. (13), and the solving procedures are similar.

III. EXPERIMENTS AND ANALYSIS

In this section, experiments to verify the performance of our l_p-norm-based underwater image restoration algorithm are described. First, we demonstrate how the parameters ξ_d, ξ_s, and ξ_i of Eq. (11) are set. Subsequently, we compare our dehazing approach with state-of-the-art enhancement techniques in underwater scenes through both visual validation and quantitative evaluation.

As shown in Eq. (11), ξ_d, ξ_s and ξ_i modulate (including stretch and compress) the detail level, structure level, and illuminance level, respectively, according to different attenuations in various underwater scenes. We apply two typical degraded underwater images that are rich in texture and structure information, and they are exhibited in two rows in Fig.(6). Column (a) shows these two original underwater images whose visibility requires improvement. In column (c), the value of the detail level parameter ξ_d is assigned with 2.0 and ξ_s, ξ_i with 0.4 and 0.7, respectively. As shown, the textures and details of the two outputs are overenhanced.

For column (d), the structure level ξ_s is amplified by almost four times to 1.5, and ξ_d, ξ_i are equal to 1.0 and 0.7, respectively. The restored images are vastly indistinct, and numerous details and textures are lost. In column (e), the illuminance level parameter ξ_i is set to 0.3, and ξ_d, ξ_s are 1.0 and 0.7, respectively. The derived results indicate apparently dark and obscure images.

Finally, we set the values of ξ_d, ξ_s, ξ_i to (1.0, 0.4, 0.7), and the restored outputs are shown in (b). As shown, the images of (b) demonstrate better perceptual quality compared to (a). Thus, the following validations are based on ξ_d, ξ_s, ξ_i of (1.0, 0.4, 0.7), respectively.

To assess our enhancing technique more accurately, we performed a quantitative evaluation based on visual validation. We implemented indicators such as the mean gradient (MG), standard deviation (SD), and contrast ratio (CR) [52] in Tab. (1). The index of the MG measures the greyscale changing ratio on the edges and small details, and it is used to characterize the clarity of the image. The larger the MG, the higher the image clarity. Conversely, the less the image hierarchy and the more blurred the image. The SD denotes the greyscale discreteness of the image. The larger its value is, the more vivid the image colour. Furthermore, the CR depicts varying grades from completely black to completely white. The larger it is, the more the grade changes from black to white occur, and consequently, the more colourful and delicate the image appears.

In Fig.(7), (b) shows the results of the outdoor dehazing approaches of He *et al.* [25]. This method performs poorly for underwater scenes, especially for hue rendition. As mentioned previously, although light propagation in hazy air and in degraded underwater scenes seems very similar, the underwater dehazing problem is more challenging. Owing to the absorption effect for underwater light, deeper seas induce more shorter-wavelength light (green and blue) casts than longer-wavelength light (red). Thus, an underwater image
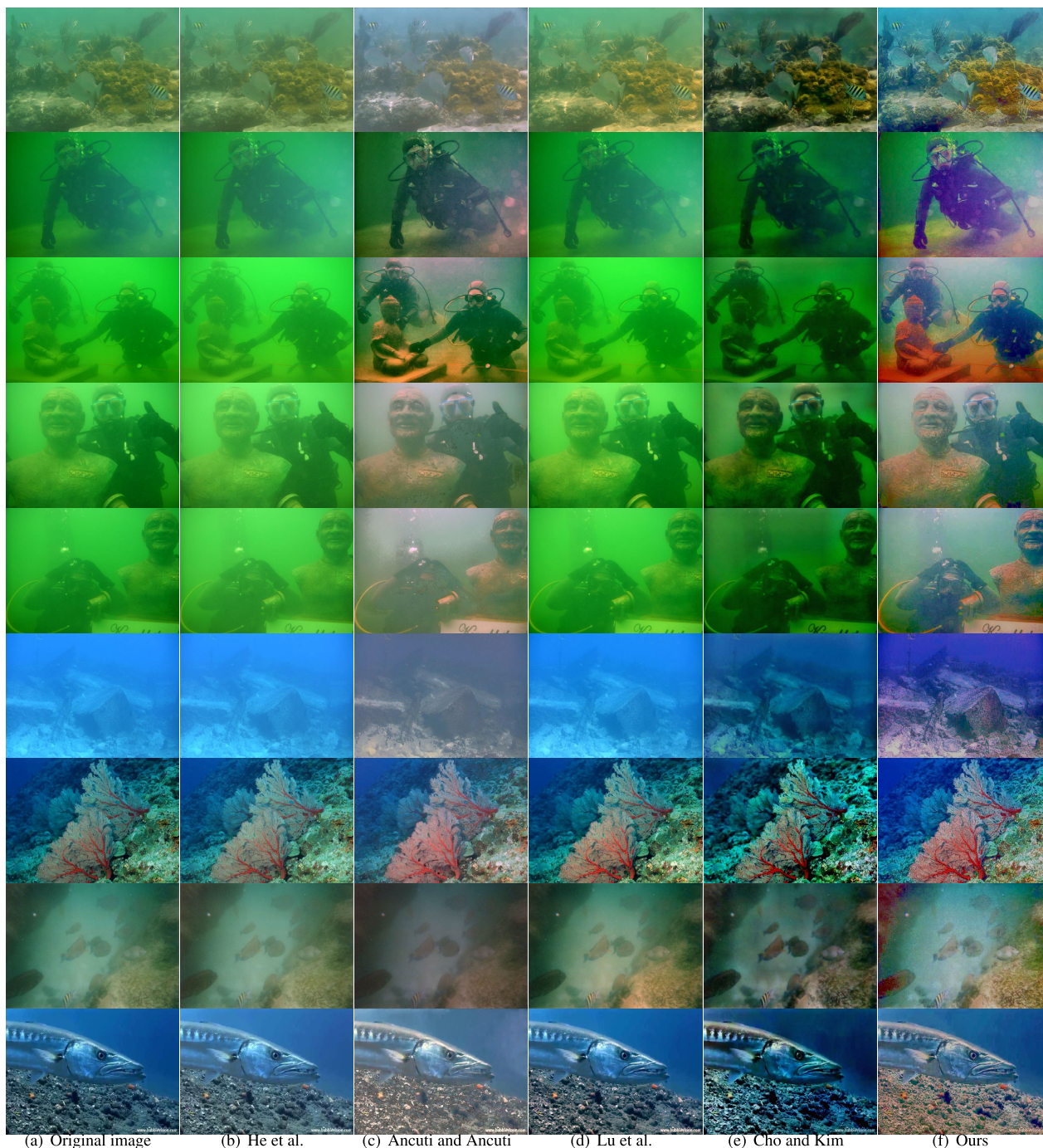
FIGURE 7. Comparison of recent dehazing techniques. Apart from the original hazy images in this figure, the results of He *et al.* [25], Ancuti *et al.* [5], Huimin *et al.* [51], Cho and Kim [50], and our technique are shown.

generally appears greenish or blueish, rendering it more difficult to restore using outdoor dehazing techniques. Tab. (1) shows that indicators of the MG, SD, and CR in (b) are generally lower compared with underwater specialized enhancement methods of Ancuti, Lu and Cho. The discrepancies are more obvious compared with our proposed underwater specialized method. Fig.(7) also presents similar results.

For the underwater specialized algorithms of Ancuti *et al.* [5], Huimin *et al.* [51], and Cho and Kim [50],

shown in (c), (d) and (e). Ancuti et al.'s approach is the multi-scale fusion method based on a Laplacian pyramid. Lu et al.'s approach compensates for the underwater attenuation discrepancy and enhances the distorted colour in turbid water in shallow regions. Cho et al. combined a model-based and multi-band fusion-based method and presented colour-corrected image enhancement while elaborating image details. They generally exhibit better performance than the method of He *et al.* [25] on both visual perception

TABLE 1. Qualitative comparison of the nine corresponding images that are shown in Fig.(7) (with same order). MG represents the mean gradient, SD represents the standard deviation, and CR represents the contrast ratio. A larger metric is better.

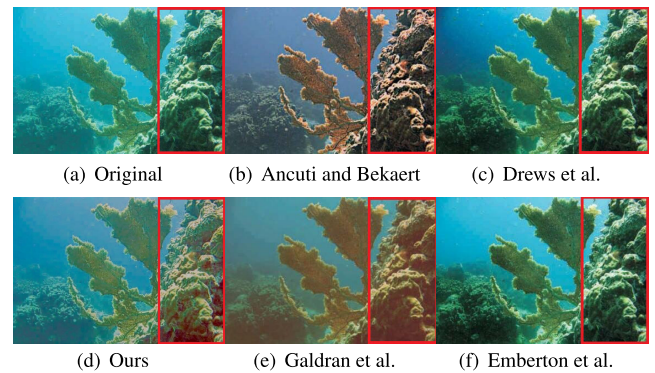|	Original image			He			Ancuti			Lu			Cho and Kim			Ours		
	MG	SD	CR	MG	SD	CR	MG	SD	CR	MG	SD	CR	MG	SD	CR	MG	SD	CR
1	0.0076	0.0794	2.5658	0.0070	0.0746	2.5480	0.0089	0.0821	2.4835	0.0091	0.0889	2.5605	0.0126	0.1233	2.9164	0.0192	0.1223	2.7963
2	0.0027	0.0749	2.5179	0.0025	0.0732	2.5265	0.0061	0.0918	2.6015	0.0032	0.0854	2.5425	0.0031	0.0592	2.5242	0.0091	0.1463	3.2839
3	0.0033	0.0925	2.5476	0.0031	0.0901	2.5470	0.0097	0.1427	2.9908	0.0039	0.0967	2.5813	0.0040	0.0898	2.5594	0.0098	0.1625	3.4175
4	0.0042	0.1333	2.9487	0.0039	0.1281	2.8914	0.0053	0.1234	2.7624	0.0049	0.1455	3.1049	0.0052	0.1343	2.8699	0.0089	0.1511	3.2930
5	0.0035	0.1005	2.5787	0.0032	0.0972	2.5649	0.0057	0.1083	2.5656	0.0041	0.1085	2.6147	0.0041	0.0943	2.5697	0.0091	0.1359	2.9409
6	0.0054	0.0588	2.1757	0.0050	0.0555	2.1813	0.0076	0.0725	2.0811	0.0065	0.0719	2.1286	0.0075	0.0736	2.3636	0.0201	0.1352	2.7011
7	0.0205	0.1321	2.9535	0.0200	0.1395	3.0732	0.0196	0.1306	2.8256	0.0240	0.1571	3.2983	0.0329	0.1746	3.4477	0.0320	0.1610	3.3571
8	0.0039	0.1147	2.6778	0.0037	0.1083	2.6193	0.0034	0.0756	2.5275	0.0044	0.1367	2.9716	0.0058	0.1055	2.7204	0.0106	0.1299	2.9490
9	0.0267	0.1230	2.5341	0.0259	0.1196	2.5770	0.0318	0.1227	2.2536	0.0309	0.1410	2.8513	0.0374	0.1747	3.2058	0.0322	0.1384	3.0269

TABLE 2. Process of underwater image enhancement based on L_p-norm decomposition.

1: **INPUT:** Original image
2: **OUTPUT:** Enhanced image
3:
4: Initialize the parameters $(\xi_d, \ \xi_s, \ \xi_i)$.
5:
6: % **FIRST INPUT CHANNEL**
7:    Input1 ← Original image;
     %% Colour correction using histogram normalization colour balance method.
8:    $(\xi_d D_t^1, \ \xi_s S_t^1, \ \xi_i I_l^1)$ ← Input1;
     %% Decomposition of input1 based on l_p-norm.
     %% S_t^1, I_l^1 are of 1/2 and 1/4 downsampling respectively.
9:
10: % **SECOND INPUT CHANNEL**
11: Input2 ← Input1;   % Contrast enhance using contrast local adaptive histogram equalization.
12: $(\xi_d D_t^2, \ \xi_s S_t^2, \ \xi_i I_l^2)$ ← Input2;
     %% Decomposition of input2 based on l_p-norm.
     %% S_t^2, I_l^2 are of 1/2 and 1/4 downsampling respectively.
13:
14: % **WEIGHT MAPS**
15: $\bar{W}^1$ ← Input1;
     $\bar{W}^2$ ← Input2;
     %% Calculate normalized weight $\bar{W}^1$ and $\bar{W}^2$ from Global contrast, Local contrast and Saliency weights of Input1 and Input2 respectively.
16: Weight1 (lev.1,lev.2,lev.3)← $\bar{W}^1$;
     Weight2 (lev.1,lev.2,lev.3)← $\bar{W}^2$;
     %% Weight1 and Weight2 are three-level Gaussian pyramids derived by Gaussian filter from $\bar{W}^1$ and $\bar{W}^2$ respectively.
     %% lev.1, lev.2, lev.3 denote the three levels respectively.
17:
18: % **FUSION**
19: Enhanced image=
     $\xi_d D_t^1 \cdot Weight1(lev.1) + \xi_d D_t^2 \cdot Weight2(lev.1)$
     $+\{\xi_s S_t^1 \cdot Weight1(lev.2) + \xi_s S_t^2 \cdot Weight2(lev.2)\}\uparrow^2$
     $+\{\xi_i I_l^1 \cdot Weight1(lev.3) + \xi_i I_l^2 \cdot Weight2(lev.3)\}\uparrow^4$.
     %% Pyramid reconstruct and fusion.
     %% $\uparrow^2$ and $\uparrow^4$ are of 2 and 4 upsampling respectively.
end

(a) Original   (b) Ancuti and Bekaert   (c) Drews et al.

(d) Ours   (e) Galdran et al.   (f) Emberton et al.

FIGURE 8. Comparison of different dehazing approaches (Ancuti and Bekaert [6], Paulo Drews *et al.* [32], our approach, Galdran *et al.* [24], and Emberton *et al.* [53]).

and quantitative evaluation. Among these three approaches, they generally have similar enhancing effects. Under further analysis, Ancuti's method presents better clarity on images with fewer detail components, such as the second, third and fifth rows, from top to bottom. However, due to the bilateral filter, which can smooth small textures and sharpen evident edges, this method shows worse clarity due to texture lost over images with more detailed components, such as the

first, seventh, eighth and ninth rows from top to bottom. This method also shows a more realistic chromatic appearance over almost all images. For these images with more detailed components, Lu and Cho's methods show higher clarity than Ancuti's methods. Moreover, over almost all images, Cho's method shows the highest contrast due to the unreal appearance of over enhancement. In general, fusion strategy-based approaches such as Ancutis and Chos are controllable and flexible; thus, they can generate better appearance regardless of qualitative or quantitative evaluations.

Compared with our proposed technique, our work outperforms those of Ancuti et al., Lu et al. and Cho and Kim in terms of clarity index MG and chromatic parameter SD and CR. It is noteworthy that some outputs of Ancuti et al. (such as the third and fourth rows from top to bottom) exhibit slightly better visual perceptions, however, with worse indicators of MG, SD and CR than ours. Through a closed inspection, we deduce that this is because of the bilateral filter. Although the bilateral filter suppresses noise, it also smooths small textures. This method improves visibility at the cost of information loss of details. For the first, seventh and ninth rows of (e) from bottom to top, the contrast index CR in Tab. (1) presents a higher value than our approach. This is because the colour over enhancement of Cho and Kim, which results in an unrealistic image of the underwater scene.

To enrich our visual assessment, Fig.(8) shows a comparison of our work with other recent methods of Ancuti and Bekaert [6], Paulo Drews *et al.* [32], Galdran *et al.* [24], and Emberton *et al.* [53]. As shown in the red rectangular box of rocks on the right side of the picture, our approach exhibits more details, especially in dark zones. This is a benefit of the
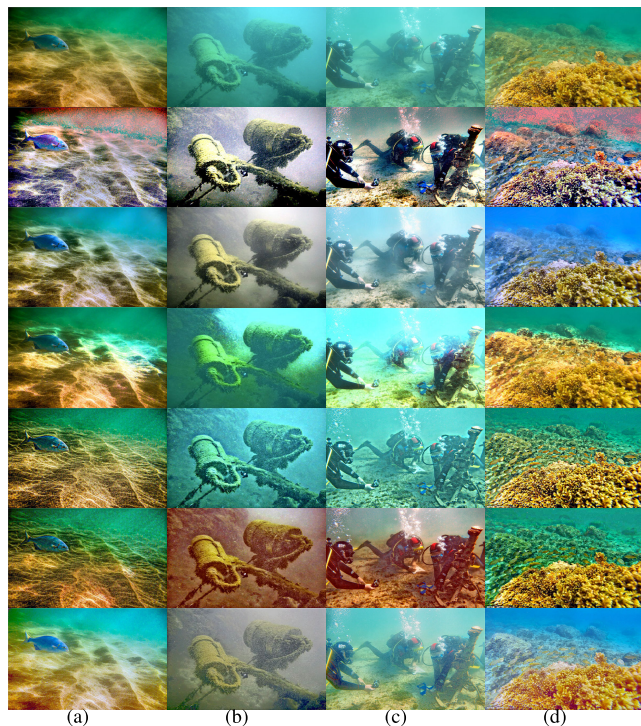
FIGURE 9. Comparison of different dehazing approaches. From the top to bottom of each column, there are original degraded input images and images processed separately by approaches of Jobson *et al.* [54], Ancuti *et al.* [5], Meng *et al.* [55], He *et al.* [25], Cho *et al.* [56], and our group.

texture and edge enhancement and the illuminant dynamic range depressing.

Fig.(9) presents the other four images derived separately from our approach and those of Jobson *et al.* [54], Ancuti *et al.* [5], Meng *et al.* [55], He *et al.* [25], and Cho *et al.* [56]. Outputs of these methods are arranged from top to bottom in each column below the first row of the original input images. Among them, the performances of our approach, Ancuti *et al.* approach, He *et al.* approach and Cho *et al.* approach are relatively better. Compared with Ancuti et al.'s method, the outputs of our approach demonstrate more delicate and smoother colours, rich small details, and higher clarity. As the outdoor dehazing method carried out on underwater scenes, the results of He et al. and Cho et al. both indicate high contrast on the details but slightly unreal expression on colour. Additionally, our technique generates a relatively better illuminant dynamic range. This makes the dark zone slightly brighter, and the strong light regions become soft; thus, more details can be shown.

In conclusion, our two-channel l_p-norm-based decomposition and fusion operator yields good image quality in terms of both perceptual validation and quantitative assessment compared with state-of-art methods. The key advantage of our method is its relatively high image clarity, mainly due to well-restored details of small edges and textures and the depressed illuminance dynamic range. However, the primary limitation of our method is the relatively weak performance caused by noise. We found it is very difficult to preserve the

small details clearly while effectively depressing noise, which commonly exist in underwater surroundings.

IV. CONCLUSION

We demonstrated an alternative method based on two-channel l_p-norm decomposition to enhance underwater images herein. To the best of our knowledge, we are the first to introduce an l_p-norm decomposition strategy into a single underwater image two-channel fusion restoration or enhancement strategy. As tested on a large set of underwater images, our approach can generate outputs with good contrast, highly accurate details, and a relatively uniform dynamic brightness range. In future research, we are planning to focus on how to suppress noise while effectively recovering detailed information about underwater images.

REFERENCES

[1] D. M. Kocak, F. R. Dalgleish, F. M. Caimi, and Y. Y. Schechner, "A focus on recent developments and trends in underwater imaging," *Mar. Technol. Soc. J.*, vol. 42, no. 1, pp. 52–67, 2008.

[2] S. Fan, E. P. J. Parrott, B. S. Y. Ung, and E. Pickwell-Macpherson, "Calibration method to improve the accuracy of THz imaging and spectroscopy in reflection geometry," *Photon. Res.*, vol. 4, no. 3, pp. A29–A35, 2016.

[3] M. Le, G. Wang, H. Zheng, J. Liu, Y. Zhou, and Z. Xu, "Underwater computational ghost imaging," *Opt. Express*, vol. 25, no. 19, pp. 22859–22868, 2017.

[4] H. Lu, Y. Li, S. Nakashima, H. Kim, and S. Serikawa, "Underwater image super-resolution by descattering and fusion," *IEEE Access*, vol. 5, pp. 670–679, 2017.

[5] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *Proc. IEEE CVPR*, Jun. 2012, pp. 81–88.

[6] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert, "Color balance and fusion for underwater image enhancement," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 379–393, Jan. 2018.

[7] B. L. McGlamery, "A computer model for underwater camera systems," *Proc. SPIE*, vol. 208, pp. 221–231, Mar. 1980.

[8] J. S. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE J. Ocean. Eng.*, vol. 15, no. 2, pp. 101–111, Apr. 1990.

[9] D. Dansereau, O. Pizarro, and S. Williams, "Linear volumetric focus for light field cameras," *ACM Trans. Graph.*, vol. 34, no. 2, pp. 1–20, 2015.

[10] K. A. Skinner and M. Johnson-Roberson, "Underwater image dehazing with a light field camera," in *Proc. IEEE CVPRW*, Jul. 2017, pp. 1775–1782.

[11] Y. Y. Schechner and Y. Averbuch, "Regularized image recovery in scattering media," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1655–1660, Sep. 2007.

[12] D.-M. He and G. G. L. Seet, "Divergent-beam Lidar imaging in turbid water," *Opt. Lasers Eng.*, vol. 41, pp. 217–231, Jan. 2004.

[13] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, "A variational framework for Retinex," *Int. J. Comput. Vis.*, vol. 52, no. 1, pp. 7–23, 2003.

[14] C. Ancuti, C. O. Ancuti, C. De Vleeschouwer, and A. C. Bovik, "Nighttime dehazing by fusion," in *Proc. IEEE ICIP*, Sep. 2016, pp. 2256–2260.

[15] H. Israël and F. Kasten, "Koschmieders theorie der horizontalen sichtweite," in *Die Sichtweite im Nebel und die Möglichkeiten ihrer künstlichen Beeinflussung*, 1959, pp. 7–10.

[16] R. T. Tan, "Visibility in bad weather from a single image," in *Proc. IEEE CVPR*, Jun. 2008, pp. 1–8.

[17] C. O. Ancuti and C. Ancuti, "Single image dehazing by multi-scale fusion," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3271–3282, Aug. 2013.

[18] K. Gu, D. Tao, J.-F. Qiao, and W. Lin, "Learning a no-reference quality assessment model of enhanced images with big data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 4, pp. 1301–1313, Apr. 2018.

[19] K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 284–297, Jan. 2016.

[20] K. Gu, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "Automatic contrast enhancement technology with saliency preservation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 9, pp. 1480–1494, Sep. 2015.

[21] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, 2008.

[22] K. Nishino, L. Kratz, and S. Lombardi, "Bayesian defogging," *Int. J. Comput. Vis.*, vol. 98, no. 3, pp. 263–278, Jul. 2012.

[23] R. Liu, L. Ma, Y. Wang, and L. Zhang, "Learning converged propagations with deep prior ensemble for image enhancement," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1528–1543, Mar. 2019.

[24] A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, "Automatic red-channel underwater image restoration," *J. Vis. Commun. Image Represent.*, vol. 26, pp. 132–145, Jan. 2015.

[25] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE CVPR*, Jun. 2009, pp. 1956–1963.

[26] K. B. Gibson and T. Q. Nguyen, "On the effectiveness of the dark channel prior for single image dehazing by approximating with minimum volume ellipsoids," in *Proc. IEEE ICASSP*, May 2011, pp. 1253–1256.

[27] C. Chen, M. N. Do, and J. Wang, "Robust image and video dehazing with visual artifact suppression via gradient residual minimization," in *Computer Vision—ECCV*. Springer, 2016.

[28] S.-C. Pei and T.-Y. Lee, "Nighttime haze removal using color transfer pre-processing and dark channel prior," in *Proc. IEEE ICIP*, Sep./Oct. 2012, pp. 957–960.

[29] F. Chen, S. Zhuo, X. Zhang, L. Shen, and S. Süsstrunk, "Near-infrared guided color image dehazing," in *Proc. IEEE ICIP*, Sep. 2013, pp. 2363–2367.

[30] S.-C. Huang, B.-H. Chen, and W.-J. Wang, "Visibility restoration of single hazy images captured in real-world weather conditions," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 10, pp. 1814–1824, Oct. 2014.

[31] J. Y. Chiang and Y.-C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1756–1769, Apr. 2012.

[32] P. Drews, Jr., E. do Nascimento, F. Moraes, S. Botelho, and M. Campos, "Transmission estimation in underwater single images," in *Proc. IEEE ICCVW*, Dec. 2013, pp. 825–830.

[33] Y. Zhu, Y. Zhang, and A. L. Yuille, "Single image super-resolution using deformable patches," in *Proc. IEEE CVPR*, Jun. 2014, pp. 2917–2924.

[34] X. Gao, K. Zhang, D. Tao, and X. Li, "Image super-resolution with sparse neighbor embedding," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3194–3205, Jul. 2012.

[35] X. Ye and A. Yuille, "Learning a dictionary of deformable patches using GPUs," in *Proc. IEEE ICCVW*, Nov. 2011, pp. 483–490.

[36] X. Gao, K. Zhang, D. Tao, and X. Li, "Image super-resolution with sparse neighbor embedding," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3194–3205, 2012.

[37] Z. Zhang, Z. Wang, Z. Lin, and H. Qi, "Image super-resolution by neural texture transfer," in *Proc. IEEE CVPR*, Jun. 2019, pp. 7982–7991.

[38] Z. Murez, T. Treibitz, R. Ramamoorthi, and D. J. Kriegman, "Photometric stereo in a scattering medium," in *Proc. IEEE ICCV*, Dec. 2015, pp. 3415–3423.

[39] C. Tsiotsios, M. E. Angelopoulou, A. J. Davison, and T.-K. Kim, "Effective backscatter approximation for photometry in murky water," 2016, *arXiv:1604.08789*. [Online]. Available: https://arxiv.org/abs/1604.08789

[40] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep photo: Model-based photograph enhancement and viewing," *ACM Trans. Graph.*, vol. 27, no. 5, 2008, Art. no. 116.

[41] Z. Liang, J. Xu, D. Zhang, Z. Cao, and L. Zhang, "A hybrid l1-l0 layer decomposition model for tone mapping," in *Proc. IEEE CVPR*, Jun. 2018, pp. 4758–4766.

[42] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE CVPR*, Jun. 2009, pp. 1597–1604.

[43] T. Pu and G. Ni, "Contrast-based image fusion using the discrete wavelet transform," *Opt. Eng.*, vol. 39, no. 8, pp. 2075–2082, 2000.

[44] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. no. 67.

[45] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," *Int. J. Comput. Vis.*, vol. 81, no. 1, pp. 24–52, 2009.

[46] H. Hermann, "Art and visual perception," Tech. Rep., 1978.

[47] C. Grigorescu, N. Petkov, and M. A. Westenberg, "Contour and boundary detection improved by surround suppression of texture edges," *Image Vis. Comput.*, vol. 22, no. 8, pp. 609–622, 2004.

[48] S. Bi, X. Han, and Y. Yu, "An L_1 image transform for edge-preserving smoothing and scene-level intrinsic decomposition," *ACM Trans. Graph.*, vol. 34, no. 4, 2015, Art. no. 78.

[49] J. Chang, R. Cabezas, and J. W. Fisher, III, "Bayesian nonparametric intrinsic image decomposition," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 704–719.

[50] Y. Cho and A. Kim, "Visibility enhancement for underwater visual SLAM based on underwater light scattering model," in *Proc. ICRA*, May/Jun. 2017, pp. 710–717.

[51] L. Huimin, L. Yujie, Z. Lifeng, and S. Seiichi, "Contrast enhancement for images in turbid water," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 32, no. 5, pp. 886–893, 2015.

[52] Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai, "No-reference quality assessment of contrast-distorted images based on natural scene statistics," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 838–842, Jul. 2015.

[53] S. Emberton, L. Chittka, and A. Cavallaro, "Hierarchical rank-based veiling light estimation for underwater dehazing," in *Proc. BMVC*, 2015, pp. 1–12.

[54] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "A multiscale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 1997.

[55] G. Meng, W. Ying, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *Proc. IEEE ICCV*, Dec. 2013, pp. 617–624.

[56] Y. Cho, J. Jeong, and A. Kim, "Model-assisted multiband fusion for single image enhancement and applications to robot vision," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 2822–2829, Oct. 2018.

JIANHUA WANG received the M.S. degree in optical electromechanical integration from the Guilin University of Electronic Technology, in 2012. He is currently pursuing the Ph.D. degree with Hohai University, under the supervision of Prof. H. Wang. His research interests include computer vision, artificial intelligence, data mining, and underwater navigation.

HUIBIN WANG received the Ph.D. degree from the China University of Mining and Technology, Xuzhou, China, in 2003. He is currently a Professor with the College of Computer and Information Engineering, Hohai University, Nanjing, China. His research interests include multisensor systems and information fusion, image processing, and pattern recognition.

GUOWEI GAO received the B.S. degree in computational science and the M.S. degree in applied mathematics from Henan University, Kaifeng, China, in 2005 and 2008, respectively, and the Ph.D. degree in information and communication engineering from Hohai University, Nanjing, China, in 2018. He has been with the School of Software Engineering, Anyang Normal University, Anyang, China, since 2008.

HUIMIN LU received the B.S. degree in electronics information science and technology from Yangzhou University, in 2008, the M.S. degrees in electrical engineering from the Kyushu Institute of Technology and Yangzhou University, in 2011, and the Ph.D. degree in electrical engineering from the Kyushu Institute of Technology, in 2014, where he was a JSPS Research Fellow at the Kyushu Institute of Technology, from 2013 to 2016. He is currently an Assistant Professor with the Kyushu Institute of Technology. His research interests include computer vision, robotics, artificial intelligence, and ocean observing.

ZHEN ZHANG received the B.Sc. and Ph.D. degrees from Hohai University, in 2007 and 2013, respectively, where he is currently a Lecturer. His main research interests include optical imaging and multisensor systems, computer vision and digital image processing, and image-based flow measurement.

● ● ●