Human Visual System based Automatic Underwater Image Enhancement in NSCT domain

Yan Zhou, Qingwu Li, Guanying Huo
Key Laboratory of Sensor Networks and Environmental Sensing, Hohai University, Changzhou 213022, China
[e-mail: strangeryan@163.com, liqw@hhuc.edu.cn, huoguanying@163.com]
*Corresponding author: Qingwu Li

Received November 17, 2014; revised May 27, 2015; accepted December 23, 2015; published February 29, 2016

Abstract

Underwater image enhancement has received considerable attention in last decades, due to the nature of poor visibility and low contrast of underwater images. In this paper, we propose a new automatic underwater image enhancement algorithm, which combines nonsubsampled contourlet transform (NSCT) domain enhancement techniques with the mechanism of the human visual system (HVS). We apply the multiscale retinex algorithm based on the HVS into NSCT domain in order to eliminate the non-uniform illumination, and adopt the threshold denoising technique to suppress underwater noise. Our proposed algorithm incorporates the luminance masking and contrast masking characteristics of the HVS into NSCT domain to yield the new HVS-based NSCT. Moreover, we define two nonlinear mapping functions. The first one is used to manipulate the HVS-based NSCT contrast coefficients to enhance the edges. The second one is a gain function which modifies the lowpass subband coefficients to adjust the global dynamic range. As a result, our algorithm can achieve contrast enhancement, image denoising and edge sharpening automatically and simultaneously. Experimental results illustrate that our proposed algorithm has better enhancement performance than state-of-the-art algorithms both in subjective evaluation and quantitative assessment. In addition, our algorithm can automatically achieve underwater image enhancement without any parameter tuning.

Keywords: Underwater image, automatic enhancement, image denoising, human visual system, nonsubsampled contourlet transform

http://dx.doi.org/10.3837/tiis.2016.02.022 **ISSN : 1976-7277**

1. Introduction

Underwater optical imaging is widely used in practical applications such as underwater archaeology, water conservancy project, marine biological monitoring and geological or biological environments assessment. Acquiring clear images from underwater environments is an important issue in these scientific missions [1]. However, underwater optical images are notorious for poor visibility, low contrast, edge-blurring and being full of noise, owing to the weak and non-uniform illumination, light absorption and light scattering caused by suspended particles in water, etc. Therefore it is necessary to improve the visual quality of underwater images for further image processing, such as edge detection, image segmentation, target recognition and so on.

Generally, image enhancement approaches can be divided into two categories: spatial domain methods and transform domain ones. Spatial domain methods deal with the image pixels to achieve desired enhancement. Such methods include histogram equalization (HE), linear contrast stretch, convolution mask enhancement, etc. Many of these methods simply define a global or adaptive one-to-one mapping by which the intensity values of individual pixels are modified [2]. These approaches often yield inadequate detail preservation or over-enhancement as some local details and noise in the image are not considered. For example, HE has received considerable attention due to its simple and straightforward implementation. It uses an input-to-output mapping obtained from the cumulative distribution function (CDF), which is the integral of the image histogram [3]. Although HE can efficiently utilize display intensities to achieve global image enhancement, it tends to over-enhance the image contrast if there are high peaks in the histogram, often resulting in a harsh and noisy appearance of the output image [4]. The multiscale retinex (MSR) [5] algorithm is another type of spatial domain method which is developed to attain lightness and color constancy for machine vision applications. The constancy refers to the resilience of perceived color and lightness to spatial and spectral illumination variations. Retinex theory is a powerful tool that can explain many of the basic phenomena of human visual perception and is consistent with the psycho-visual laws of the human vision system (HVS) [6]. The benefits of MSR algorithm include dynamic range compression and color independence from the spatial distribution of the scene illumination. However, this algorithm can result in "halo" artifacts, particularly in boundaries between large uniform regions. Moreover, a "graying out" can occur, i.e., the scene tends to become middle gray.

Transform domain methods often decompose the image into different frequency bands and process the image in each subband independently by using such transforms as Fourier transform (FT), discrete cosine transform (DCT), wavelet transform (WT), etc. Two-dimensional (2-D) WT commonly used is a separable extension of 1-D WT, which does not work well in capturing the image's geometric edges because of its isotropy. Multiscale geometric analysis (MGA) is developed to overcome the weakness of the separable WT in sparsely representing lines, curves and edges. In 2003, M. N. Do and M. Vetterli proposed the contourlet transform (CT) which possesses very high directional sensitivity and anisotropy [7,8]. Because of its shift-variance property, CT will cause pseudo-Gibbs phenomena around singularities. Then J. P. Zhou *et al.* proposed nonsubsampled contourlet transform (NSCT) which is a shift-invariant version of CT [9,10]. Due to the properties of directionality, anisotropy and shift-invariance, NSCT has been widely applied in image enhancement, segmentation, edge detection and so on [11-14]. In [15], H. Soyel *et al.* proposed a

NSCT-based automatic image enhancement method which enhances the contrast by modifying the NSCT coefficients with a nonlinear mapping function. This state-of-the-art image enhancement method has achieved good results in both grayscale and colour images. However, since it does not take into account of the physical properties in underwater environments, for some underwater images, this method could not sufficiently adjust contrast, suppress noise and amplify subtle details in some local regions to correct non-uniform illumination.

Many existing multiscale transform domain enhancement algorithms achieve contrast enhancement by directly manipulating the transform coefficients [16], in which case the contrast is defined in terms of absolute luminance changes. However, the HVS is sensitive to relative luminance changes rather than absolute luminance changes, which is known as the luminance masking (LM) feature of the HVS [17]. So the transform coefficients integrating this masking effect are more consistent with the HVS than the transform coefficients alone. Another important feature of the HVS is that it is sensitive to relative changes in contrast. In other words, the visibility of a stimulus is affected when surrounding stimuli exist, which is known as the contrast masking (CM) feature of the HVS [17]. The image processing algorithms which consider both the LM and CM features of the HVS have displayed the dramatic performance in edge detection and compression applications [17-19]. However, the kind of methodology which can incorporate the Retinex, LM and CM features of the HVS into NSCT domain for automatic underwater image enhancement and denoising, has not been explored.

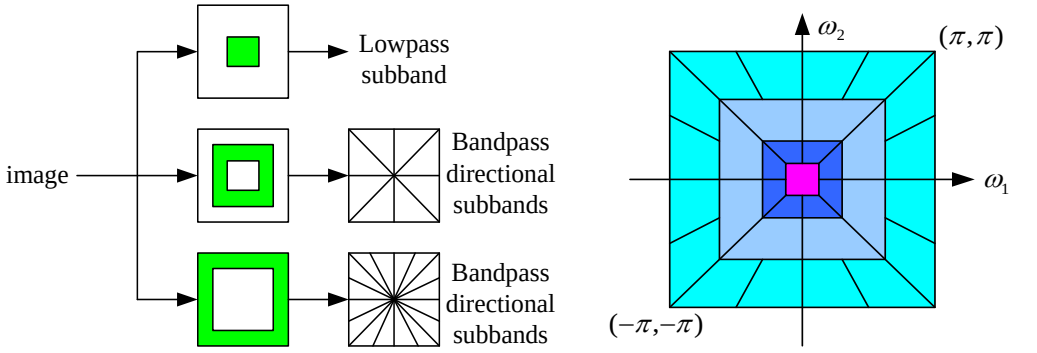
In this paper, a novel HVS-based automatic underwater image enhancement method in NSCT domain is proposed. Firstly, MSR approach based on the HVS is introduced into NSCT domain in order to eliminate non-uniform illumination, and threshold denoising is also adopted to reduce noise. Secondly, the novel HVS-based NSCT is yielded by integrating the LM and CM characteristics of the HVS into NSCT. Lastly, a new nonlinear mapping function and a new nonlinear gain function, both of which are fully automatic without any parameter tuning, are designed to manipulate the HVS-based NSCT contrast coefficients and the NSCT lowpass subband coefficients respectively. Experimental results show that our proposed method can effectively enhance the contrast, emphasize weak edges and suppress noise, producing better enhanced results than existing methods. Moreover, this method is free of parameters adjusting and can be effectively applied in machine vision systems for underwater image processing.

The rest of this paper is organized as follows: Section 2 describes the nonsubsampled contourlet transform. Section 3 analyzes the HVS perceptional properties which will be used in our algorithm. Section 4 introduces the new HVS-based NSCT and presents an automatic underwater image enhancement method which incorporates MSR, LM and CM features of the HVS into NSCT domain. The experimental results and performance evaluation are given in Section 5. The conclusions are drawn finally in Section 6.

2. Nonsubsampled Contourlet Transform

Considering the limitation of separable wavelet transform in capturing directional information, researchers have recently developed multiscale and multidirectional representations, e.g., contourlet transform, which can capture the intrinsic geometrical structures such as smooth contours in natural images [7]. A drawback of contourlet transform is its shift-variant due to downsampling and upsampling [8]. To overcome this shortcoming, A. L. Cunha *et al.* proposed an overcomplete transform called NSCT in 2006 [10]. The analysis part of NSCT is

shown in **Fig. 1**. **Fig. 1(a)** displays the nonsubsampled filter bank (NSFB) structure of NSCT. NSFB can be divided into two shift-invariant parts: 1) nonsubsampled pyramid (NSP) that provides multiscale decomposition and 2) nonsubsampled directional filter bank (NSDFB) that provides directional decomposition. First, a NSP splits the input into a lowpass subband and a highpass subband. Then a NSDFB decomposes the highpass subband into several directional subbands. The scheme is iterated repeatedly on the lowpass subband outputs of NSP. The 2-D frequency plane in the subbands split by the structure is illustrated in **Fig. 1(b)**.



(a) NSCT filter bank structure (b) NSCT frequency domain partitioning

Fig. 1. Nonsubsampled Contourlet Transform

The multiscale property of the NSCT is obtained from a shift-invariant filtering structure of NSP. NSP is achieved by using two-channel nonsubsampled 2-D filter banks. The multidirection property of the NSCT is obtained from a shift-invariant filtering structure of NSDFB, which is constructed by eliminating the downsamplers and upsamplers in the DFB. This can be achieved by first switching off the downsamplers/upsamplers in each two-channel filter bank in the DFB tree structure and then upsampling the filters accordingly.

By using NSP and NSDFB, NSCT can effectively capture geometry and directional information of images. Comparing with contourlet and other multiscale transforms, NSCT has the property of shift invariance. Therefore, NSCT can provide richer time information and more accurate frequency localised information. Furthermore, since each pixel of the transform subbands corresponds to that of the original image in the same spatial location, we can gather the geometrical information pixel by pixel from the NSCT coefficients. Based on the observations above, we use NSCT as the multiscale transform tool to achieve underwater image enhancement.

3. Human Visual Perception Characteristics

3.1 Image Representation Mechanism of the HVS

Human visual information processing is multilevel and very complex, and human visual perception is closely related to retina and visual cortex. There are two classes of light receptors: cones and rods over the surface of the retina. Cones are able to resolve fine details because each cone is connected to its own nerve end, while rods serve to give a general, overall picture of the field of view because several rods are connected to a single nerve end. As the cones and rods capture details at different resolutions simultaneously, it can be said that the retina decomposes the image into many resolutions simultaneously.

In addition, it is well-known that the receptive fields of simple cells in mammalian primary visual cortex can be characterized as being spatially localized, oriented, and bandpass [20]. Since the cells in the primary visual cortex possess a multiresolution nature, the perception of the HVS to a certain field of view can be described by using a series of gradually smoothed images and the differences between successively smoothed images [21]. The HVS has been tuned so as to capture the essential information of a natural scene using a least number of visual active cells, which is called as a sparse code on a natural scene. This result suggests that for an efficient computational image representation, it should be based on a local, directional, and multiresolution expansion. From the above discussions, the principle of NSCT corresponds to the mechanism of human visual perception.

3.2 Multiscale Retinex

The retinal-cortical theory (Retinex theory) was conceived by E. H. Land *et al.* as a model of the lightness and color perception of human vision [22]. Sensations of color show a strong correlation with reflectance, even though the amount of visible light reaching the eye depends on the product of reflectance and illumination. Human vision computes lightness and color so as to be relatively independent of spectral variations in illumination, in another word, it is lightness and color constancy. In addition, the dynamic range compression of Retinex can eliminate shadows and lighting variations in images. According to the principle of Retinex theory, an image can be defined as follows:

$$I(x,y) = L(x,y)R(x,y) \tag{1}$$

where $I(x,y)$ is the amount of light perceived by human eyes, $L(x,y)$ represents the illumination, and $R(x,y)$ is the reflectance which contains the object's characteristics.

D. J. Jobson *et al.* developed the multiscale retinex (MSR) approach which applies the center/surround retinex model for light and color rendition as well as dynamic range compression [5]. It works by processing the image with a high-pass filter, using several different widths and averaging the results together. Mathematically, the MSR output of a grayscale image can be expressed as follows:

$$R^{MSR}(x,y) = \sum_{n=1}^{N} \omega_n \left\{ \log I(x,y) - \log[F_n(x,y) * I(x,y)] \right\} \tag{2}$$

where $R^{MSR}(x,y)$ is the MSR output value, $I(x,y)$ is the input image pixel value, N is the number of scales, ω_n is the weight associated with the nth scale, "$*$" denotes the convolution operation, and $F_n(x,y)$ represents the Gaussian surround function at the nth scale as given by

$$F_n(x,y) = Ke^{-(x^2+y^2)/c_n^2} \tag{3}$$

where c_n is the Gaussian surround space constant associated with the nth scale. K is selected as

$$\iint F(x,y)dxdy = 1 \tag{4}$$

Experiments showed that a combination of three different scales (c_1 = 15, c_2 = 80, c_3=250) with equal weighting ($\omega_1 = \omega_2 = \omega_3 = 1/3$) is sufficient to provide both dynamic range compression and tonal rendition for most images [5, 23]. In this paper, we choose the parameters as listed above for MSR.

3.3 Luminance Masking

The response of HVS depends much less on the absolute luminance than on the relation of its local variations to the surrounding luminance. The minimum change required for the HVS to perceive contrast is a function of background illumination and can be approximated with four regions, namely the dark, Devries-Rose, Weber, and saturation regions, as shown in **Fig. 2** [24]. In general, the just-noticeable difference (JND) contrast threshold ΔB_t is given by

$$\Delta B_t = \frac{\Delta B}{B^{\alpha}} \tag{5}$$

where B is a background illumination and ΔB is the absolute luminance difference. In the dark region, $\alpha = 0$, indicating that the perceived contrast doesn't depend on the background luminance and is a function only of the absolute luminance difference ΔB. The value of α for the Devries-Rose region, Weber region, and saturation region as shown in **Fig. 2** are given as $\alpha = 0.5$, $\alpha = 1$, $\alpha = 2$, respectively. Generally, the denominator in Eq. (5) is calculated by some sort of average, and the numerator by some sort of difference. Weber's Contrast Law quantifies the minimum change required for a properly illuminated area which covers a very large range of background intensities. The Weber contrast of an image I is defined as

$$C_{Weber} = \frac{I - B_0}{B_0} \tag{6}$$

where B_0 indicates the background luminance obtained by uniform averaging or Gaussian filtering, and the numerator denotes the detail information.

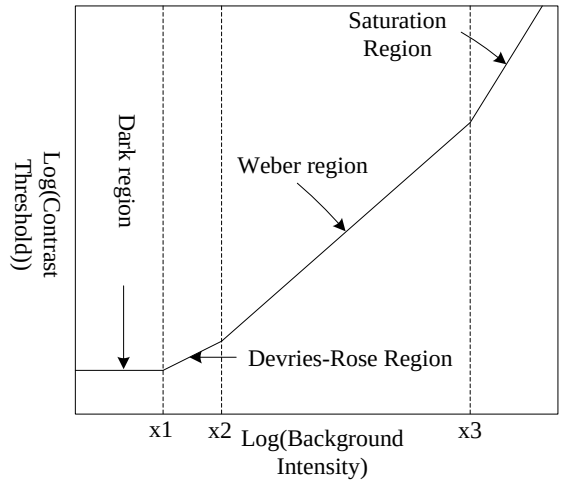


Fig. 2. Four regions of human visual response

3.4 Contrast Masking

The HVS is sensitive not only to relative changes in luminance, but also to relative changes in contrast. CM is an important phenomenon of the HVS and refers to the visibility reduction of one visual signal at the presence of another one. It means that the perceived contrast of a certain pattern that is surrounded or superimposed by other stimuli, becomes lower than its counterpart when this pattern stays alone without other stimuli. In general, the CM effect is a function of spatial frequency. The signal threshold elevation is maximal when the masker and signal have the same frequency. Threshold elevation decreases regularly as the masking frequency departs from the signal frequency [25]. Denoting signal contrast threshold by C_t , masking contrast by C, and taking the power function exponent to be 0.62, the contrast masking function is given in [25]:

$$C_t = [k(f)C]^{0.62} \tag{7}$$

where $k(f)$ is a frequency-dependent scaling factor.

4. Automatic Underwater Image Enhancement in HVS-based NSCT domain

In this section, the proposed HVS-based automatic algorithm for underwater image enhancement in NSCT domain is formulated. Firstly, the new HVS-based NSCT is presented which integrates the HVS masking properties into NSCT so as to generate a new HVS-based multiscale transform. Subsequently, the whole scheme of the automatic underwater image enhancement algorithm using the HVS-based NSCT is introduced.

4.1 HVS-based NSCT

According to Section 3, the novel HVS-based NSCT is yielded by integrating the HVS masking characteristics into NSCT domain. Since each NSCT subband coefficient corresponds to the pixel of the original image in the same spatial location, each transform subband has the same size with the original image. To obtain HVS-based NSCT, two steps are to be conducted. Firstly, similar to Weber contrast measure shown in Eq. (6), the LM contrast in NSCT domain is measured by

$$C_{LM(s,d)} = \frac{y_{(s,d)}}{\left|y_{(s,0)}\right| + c} \tag{8}$$

where $y_{(s,d)}$ is the original NSCT bandpass directional subband indexed by scale s and direction d, which denotes the high frequency details of an image. $y_{(s,0)}$ is the original NSCT lowpass subband at the sth scale, which denotes the low frequency component of an image, approximately corresponding to the background luminance. In an N level NSCT decomposition of an image, $s=N$ denotes the scale after performing NSCT decomposition procedure one time (i.e., the finest scale), and $s=1$ denotes the scale after performing NSCT decomposition procedure N times (i.e., the coarsest scale). c is a small constant equal to 0.0001 to avoid dividing by 0. $C_{LM(s,d)}$ is the output of LM contrast indexed by scale s and direction d. Secondly, the LM contrast is masked with a local activity measure (i.e., CM with Eq. (7)) to yield the LCM contrast, which is the HVS-based NSCT. The multiscale LCM contrast, which

is a function of the LM contrast, is defined as

$$C_{LCM(s,d)} = \frac{C_{LM(s,d)}}{\left|C_{LM(s-1,d)}\right|^{0.62} + c} \tag{9}$$

where c is a small constant equal to 0.0001 to avoid dividing by 0.

The NSCT coefficients can be recovered from the multiscale LCM contrast coefficients. The LM contrast can be calculated from the LCM contrast by

$$C_{LM(s,d)} = C_{LCM(s,d)} * \left(\left|C_{LM(s-1,d)}\right|^{0.62} + c\right) \tag{10}$$

and the bandpass directional subband coefficients of NSCT are calculated from the LM contrast by

$$y_{(s,d)} = C_{LM(s,d)} * \left(\left|y_{(s,0)}\right| + c\right) \tag{11}$$

Then, the image can be reconstructed from NSCT coefficients. As the proposed HVS masking transform is a completely invertible procedure, the new HVS-based NSCT is an invertible and perfect reconstruction, which generates contrast coefficients at each scale and direction of decomposition based on relevant psychophysical properties. The block diagram of the proposed HVS-based NSCT transform architecture is shown in **Fig. 3**. After NSCT decomposition, the bandpass directional subband coefficients are masked first by a measure of local luminance and then by a measure of local activity to yield the HVS-based NSCT contrast coefficients. The NSCT coefficients are recovered from the HVS-based NSCT contrast coefficients via inverse HVS masking transform.

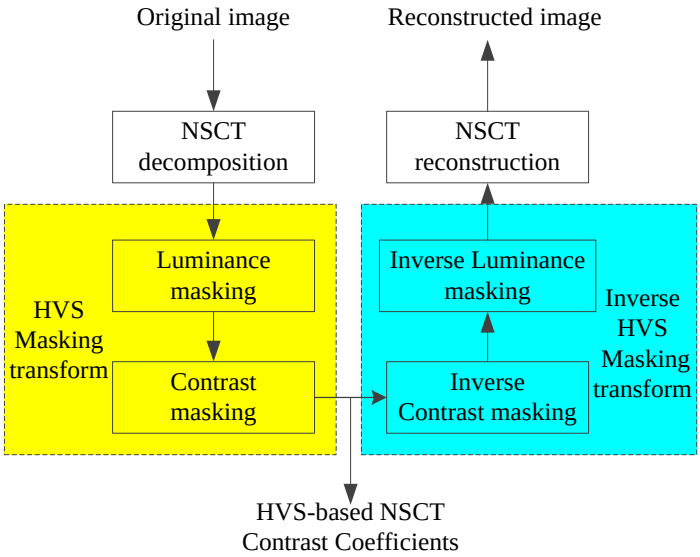


Fig. 3. Block diagram of the proposed HVS-based NSCT

4.2 Automatic Underwater Image Enhancement Algorithm

A novel automatic underwater image enhancement algorithm can be formulated using the newly proposed HVS-based NSCT. A block diagram of the proposed image enhancement algorithm is shown in **Fig. 4**. Firstly, the original image is decomposed in NSCT domain to obtain the lowpass subband coefficients and a series of the bandpass directional subband coefficients. Secondly, MSR algorithm is applied to the lowpass subband coefficients in order to remove the inhomogeneous illumination, and threshold denoising is implemented at each bandpass directional subband. Thirdly, HVS masking transform, which is shown in left dotted box of **Fig. 3**, is applied to the processed NSCT coefficients so as to yield the HVS-based NSCT contrast coefficients. A new nonlinear mapping function is designed to modify the HVS-based NSCT contrast coefficient subbands at each direction and scale of the decomposition. The overall contrast of the image is adjusted by using a new nonlinear gain function to alter the lowpass subband at the coarsest scale. Lastly, the inverse HVS masking transform, which is shown in right dotted box of **Fig. 3**, is performed to obtain the modified NSCT coefficients, and the enhanced image can be reconstructed with the modified NSCT coefficients via NSCT reconstruction.

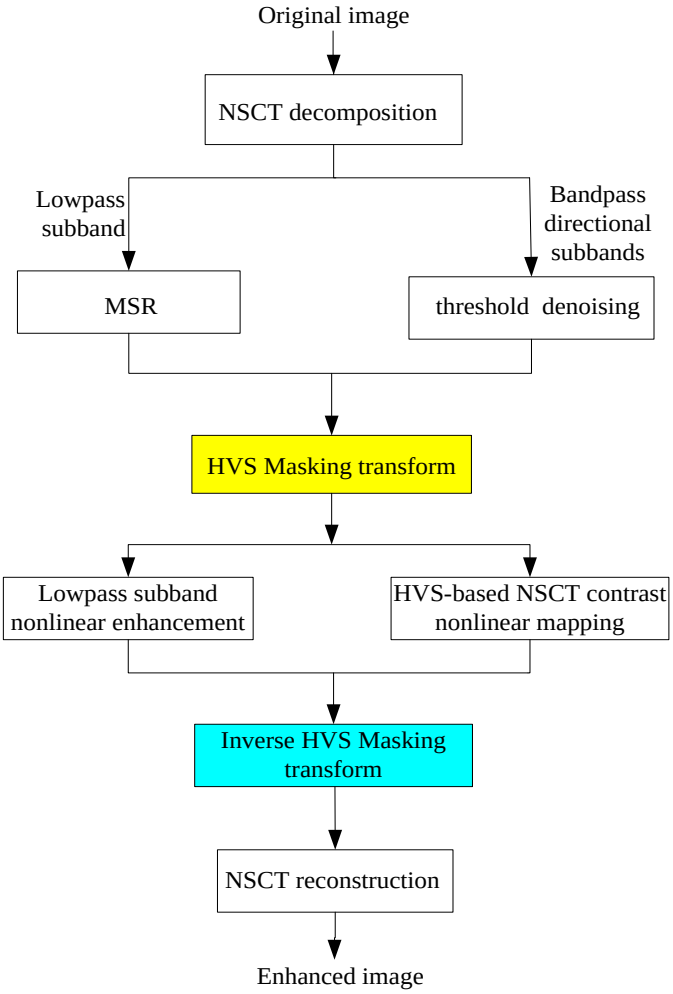


Fig. 4. Block diagram of the proposed image enhancement algorithm

After NSCT decomposition on an underwater image, the lowpass subband, which is almost noiseless, contains overall contrast information. While bandpass directional subbands contain not only edges but also noise.

Taking into account of the presence of non-uniform illumination in the underwater image, the NSCT lowpass subband is manipulated with MSR algorithm which utilizes the HVS characteristics discussed in Section 3.2. In this stage, dynamic range is properly compressed to eliminate shadows and uneven illumination.

Because edges correspond to the large NSCT coefficients and noise corresponds to the small NSCT coefficients in bandpass directional subbands, noise can be effectively suppressed by thresholding. The hard-thresholding rule is used for estimating the unknown noiseless NSCT coefficients. Thresholds for each bandpass directional subband can be chosen according to:

$$T_{s,d} = k\sigma\sqrt{\tilde{\sigma}_{s,d}} \tag{12}$$

We set k=4 for the finest scale and k=3 for the remaining ones. The noise standard deviation σ of the original image is estimated by using the robust median operator [26], i.e.,

$$\sigma = median(abs(C))/0.6745 \tag{13}$$

where C refers to the NSCT coefficients in the finest subband. An approximate value $\tilde{\sigma}_{s,d}^2$ of the individual variances at the dth directional subband of the sth scale is calculated by using Monte-Carlo simulations [27].

Subsequently, a forward HVS masking transform is applied to the NSCT coefficients to yield the HVS-based NSCT. We also propose a new nonlinear mapping function to modify the HVS-based NSCT contrast coefficients at each scale and direction independently and automatically so as to achieve multiscale contrast enhancement. The proposed nonlinear mapping function is given by:

$$\hat{C}_{LCM(s,d)} = s \cdot max(\left|C_{LCM(s,d)}\right|) \cdot sign(C_{LCM(s,d)}) \cdot \left[\sin(\frac{\pi}{2} \cdot \frac{\left|C_{LCM(s,d)}\right|}{max(\left|C_{LCM(s,d)}\right|)})\right]^{\sqrt{p}} \tag{14}$$

where

$$p = \frac{\log(\dfrac{mean(\left|C_{LCM(s,d)}\right|)}{max(\left|C_{LCM(s,d)}\right|}))}{\log[\sin(\dfrac{\pi}{2}\dfrac{mean(\left|C_{LCM(s,d)}\right|)}{max(\left|C_{LCM(s,d)}\right|}))]} \tag{15}$$

$C_{LCM(s,d)}$ is the original HVS-based NSCT contrast coefficient in the subband indexed by scale s and direction d. $\hat{C}_{LCM(s,d)}$ is the processed HVS-based NSCT contrast coefficient. $max(\left|C_{LCM(s,d)}\right|)$ denotes the maximum absolute contrast coefficient amplitude in the

subband indexed by scale s and direction d. $mean(|C_{LCM(s,d)}|)$ denotes the mean value of absolute contrast coefficient amplitude in the subband indexed by scale s and direction d. Given an N scale NSCT decomposition of an image, $C_{LCM(s,d)}$ includes N-1 scales (i.e., $1 < s \le N$), according to Eq. (9).

This nonlinear mapping function can well enhance low-contrast areas, and also avoid over-enhancement of high-contrast areas simultaneously.

Furthermore, the global dynamic range of the image can be sufficiently and accurately adjusted by using a new nonlinear gain function in the lowpass subband at the coarsest scale of NSCT decomposition. The nonlinear gain function is defined as follows:

$$\hat{y}_{(1,0)} = max(|y_{(1,0)}|) \cdot sign(y_{(1,0)}) \cdot \left[\sin\left(\frac{\pi}{2} \cdot \frac{|y_{(1,0)}|}{max(|y_{(1,0)}|)}\right) \right]^{q} \tag{16}$$

where

$$q = \frac{\log\left(\frac{mean(|y_{(1,0)}|)}{max(|y_{(1,0)}|}\right)}{\log\left[\sin\left(\frac{\pi}{2} \frac{mean(|y_{(1,0)}|)}{max(|y_{(1,0)}|}\right)\right]} \tag{17}$$

$y_{(1,0)}$ is the NSCT lowpass subband coefficient at the first scale (i.e., the coarsest scale). $\hat{y}_{(1,0)}$ is the modified NSCT lowpass subband coefficient at the first scale.

Therefore, the proposed image enhancement allows for denoising, dynamic range adjustment and contrast enhancement to be achieved automatically and simultaneously. To summarize, the proposed algorithm is described as follows:

Input: Original image I.

1) Generate an N scale NSCT decomposition of I.

2) Implement MSR algorithm in lowpass subband, and conduct threshold denoising in each bandpass directional subband.

3) Perform HVS masking transform in NSCT domain, i.e., to measure the LM contrast using Eq. (8), and to measure the LCM contrast using Eq. (9).

4) Calculate the enhanced LCM contrast using Eq. (14).

5) Calculate the enhanced NSCT lowpass subband using Eq. (16).

6) Implement the inverse HVS masking transform using Eq. (10) and Eq. (11).

7) Get the enhanced image by NSCT reconstruction.

Output: Enhanced image I'.

5. Experimental Results and Discussion

To demonstrate the effectiveness of the proposed image enhancement algorithm, we compare it with HE [3], MSR [5], and Soyel's algorithm [15]. HE and Soyel's algorithm both are fully automated image enhancement algorithms, and the MSR algorithm is implemented with

parameters listed in Section 3.2.

Image enhancement assessment is not a trivial task, and it is better to have both qualitative and quantitative assessments. Qualitative visual evaluation of the algorithms is performed using Mean Opinion Scores (MOS) [28]. The image enhancement results generated by the various algorithms are shown to a group of 30 subjects who consider themselves image processing experts, and each enhancement result is scored by assigning one of the five numeric scores (1, 2, 3, 4, and 5), where 1 indicates very bad and annoying visual quality, 5 indicates very good visual quality without annoying distortion, and other scores are selected according to the perceived image quality. These opinion scores are then averaged to yield a MOS score for each enhanced image for each algorithm. The higher the MOS score is, the better the visual quality is.

To acquire quantitative evaluation of enhancement results, the image contrast measure called the measure of enhancement by entropy, or EME using entropy is proposed in [29]:

$$EMEE = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\frac{I_{\max;m,n}}{I_{\min;m,n}+c}\ln\frac{I_{\max;m,n}}{I_{\min;m,n}+c} \tag{18}$$

where an image is split into $M\times N$ blocks $B(m,n)$ of size 4×4 , and $I_{\max;m,n}$, $I_{\min;m,n}$ are the maximum and minimum values of the pixels in each block $B(m,n)$ respectively. c is a small constant equal to 0.0001 to avoid dividing by 0.

The EME by entropy (EMEE), which is of the entropy formula form XlogX, is actually a measure of the entropy, or information, in the contrast of the image [30]. We use the EMEE here as a measure of visual quality, to compare the enhancement performance of various methods. In this case, the EMEE value should increase by a significant magnitude when the contrast of an image is enhanced noticeably, i.e., a higher EMEE value indicates an image with higher contrast.

For these experiments, our algorithm uses three scales of NSCT decomposition and four directional subbands in each scale. The test images are *Chair*, *Fish*, *Testboard*, *Archway*, *Statue and Reef* respectively with 450×327 pixels, 549×335 pixels, 256×254 pixels, 400×300 pixels, 437×328 pixels and 438×328 pixels. Illustrative examples of comparisons of the proposed approach with other existing enhancement techniques are shown in **Figs. 5-10**.

The input image *Chair* in **Fig. 5** is an underwater picture of the plastic chair in the King Cruiser wreck provided by the Bubble Vision Company. Overall, the image is dark, and in particular, the bottom half of the background is much darker than the top half of the background. HE over-enhances the image contrast with annoying amplified noise, and exaggerates the disproportionate brightness present between the top and bottom parts of the image. MSR increases the overall brightness of the image, but much of the image content tends to be a similar shade of gray, and edges become blurry. Soyel's algorithm is able to enhance the contrast in a certain extent, but the bottom of the background is still dark resulting in the loss of some subtle features therein. However, the proposed algorithm yields visually pleasing brightness, removes the non-uniformity, and enhances the weak edges while inhibiting noise efficiently.

The input image *Fish* in **Fig. 6** is extracted from an underwater footage filmed by the Bubble Vision Company. Because of light scattering under water, the fishes, the diver, and the reef at the back are hazed. HE again over-enhances the image resulting in the undesired blur of weak edges and over-brightness of top region. MSR brings a "graying out" effect of the enhanced image. Soyel's algorithm enhances the contrast of some of the image structures

slightly, but the overall contrast enhancement is not significant and fine features are still not clear. The proposed algorithm achieves both global and local enhancements simultaneously yielding the most visually pleasing output and the fine details could be enhanced sufficiently.

The input image *Testboard* in **Fig. 7** is dim in an inhomogeneous illumination. HE and Soyel's algorithm both exaggerate the non-uniformity of lighting and the local details are coverd by shadows. MSR removes shadows and lighting variations, but the enhancement of contrast is not adequate. The proposed method improves both local and global contrast, makes the illumination uniform and sharpens the faint edges obviously.

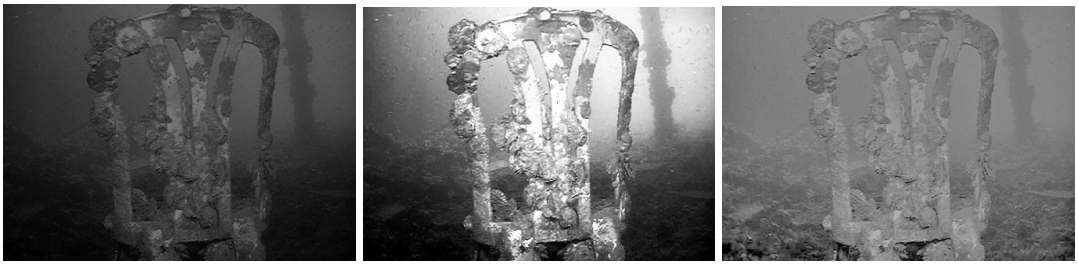
The dark input image *Archway* in **Fig. 8** shows part of the memorial archway in the flooded ancient city under Qiandao Lake, China. The proposed method is the only approach which can simultaneously enhance the contrast, sharpen the edges while reducing noise. Furthermore, it can recover the details in dark region at the bottom corner in the right.

The input image *Statue* in **Fig. 9** is dark and blurry. HE again over-enhances the image contrast and enlarges the non-uniformity of lighting. The top of the enhanced image is obviously too bright and the shadow areas appear at the right of the statue's face and body. MSR removes the lighting variations, but the scene tends to become middle gray and the details are still faint. Soyel's algorithm cannot provide a sufficient contrast enhancement. The proposed enhancement approach effectively improves the overall contrast, removes the uneven illumination and sharpens the weak edges. For example, it can be observed in **Fig. 9(e)** that the proposed algorithm significantly enhances the textures and features of the statue.

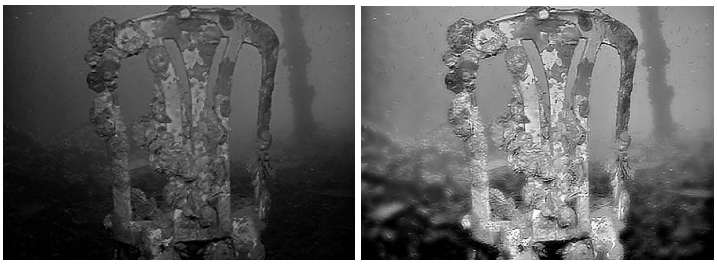
The input image *Reef* in **Fig. 10** is very obscure. It is difficult to distinguish the objects in the image. HE fails to provide an improvement due to the over-enhancement. MSR increases the brightness, but a "graying out" effect is noticeable. Soyel's algorithm produces a very slight contrast improvement. The proposed algorithm, on the other hand, produces sufficient contrast for the different objects to be recognized. It is even able to recover the details which were lost in the original image.

In all the experimental cases, MOS and EMEE measurements listed in **Table 1** and **Table 2**, respectively, agree with the subjective observations, demonstrating that the proposed algorithm outperforms other algorithms.

All the algorithms are implemented under MATLAB R2011b environment on a PC with 2.4-GHz Intel(R) Xeon(R) CPU and 8-GB RAM. The running times of the four methods are given in **Table 3**. For all the methods, the running time is proportional to the size of the image. The larger the size of the image is, the longer the running time is. The results indicate that our algorithm and Soyel's algorithm which are both based on NSCT, consume more time than other two algorithms because NSCT decomposition is time-consuming. Meanwhile, compared with Soyel's algorithm, our algorithm produces much better enhanced images though it costs a little more time. Therefore, the proposed algorithm can automatically achieve better enhancement results at the cost of a higher computational complexity which is mainly caused by NSCT decomposition.

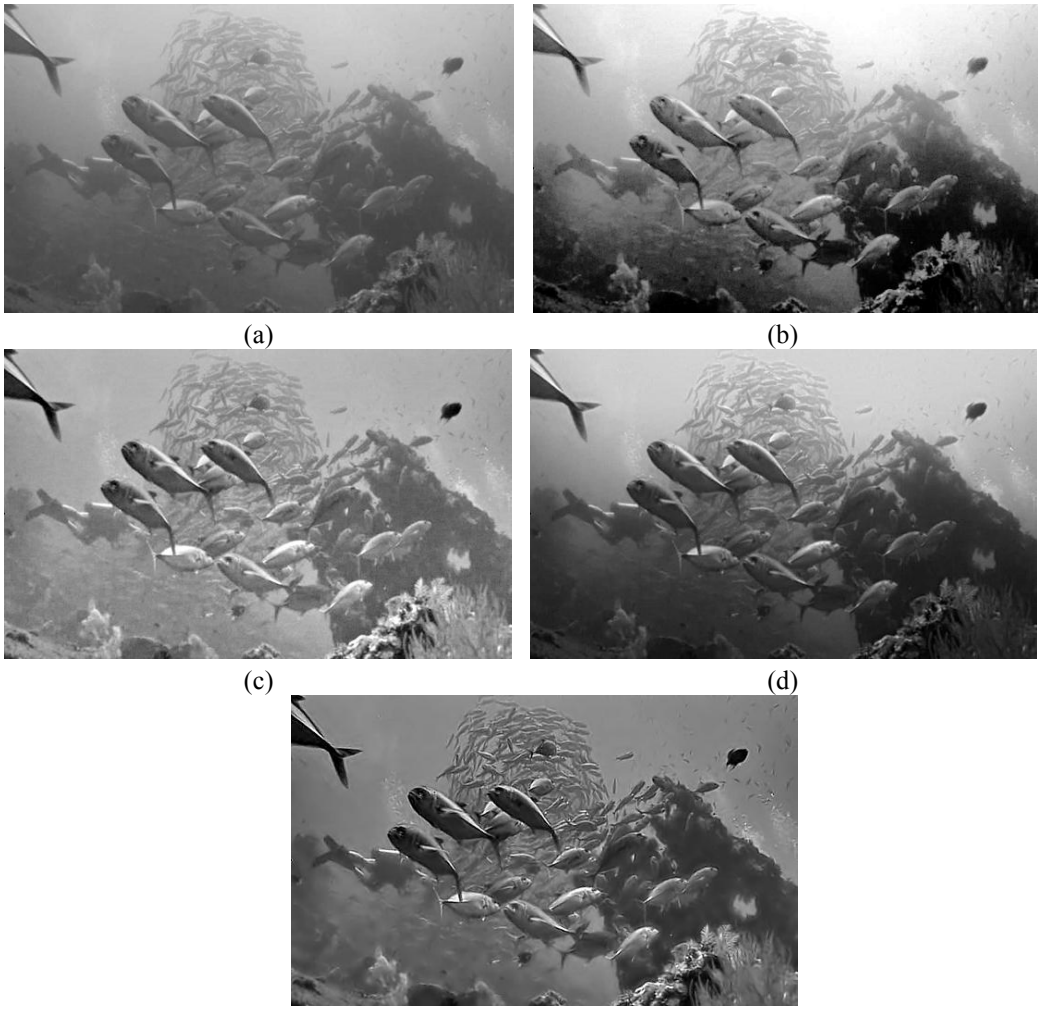


(a) (b) (c)



(d) (e)

Fig. 5. Enhancement results for image *Chair*: (a) Original image, (b) HE, (c) MSR, (d) Soyel's algorithm, (e) proposed algorithm.

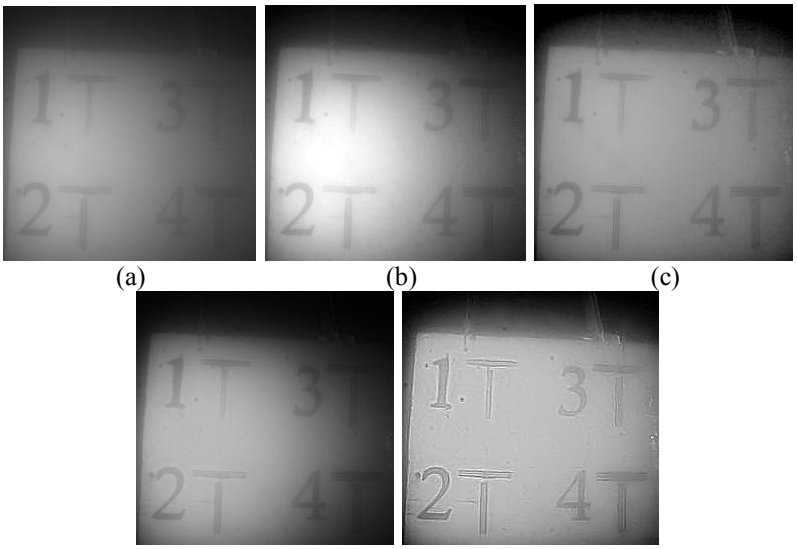


(a) (b)

(c) (d)

(e)

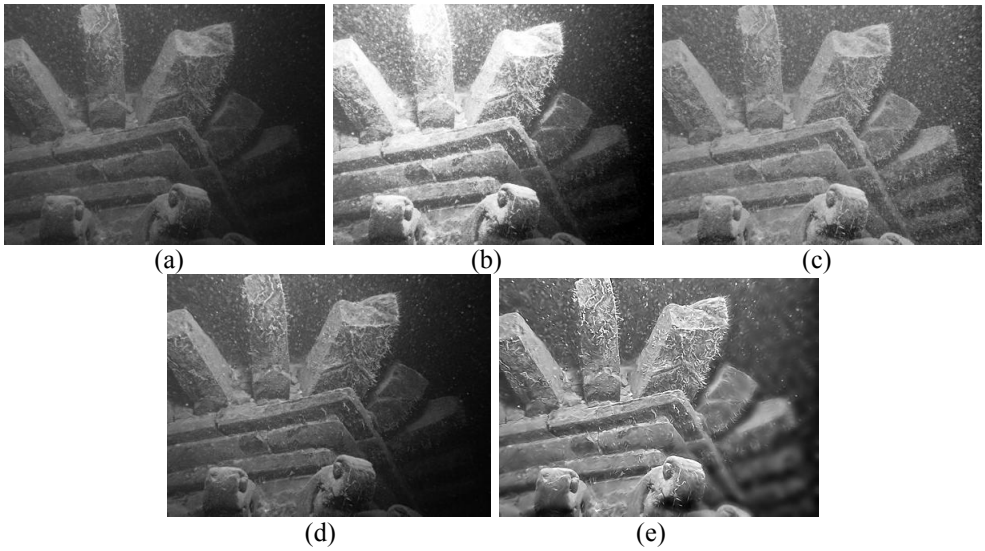
Fig. 6. Enhancement results for image *Fish*: (a) Original image, (b) HE, (c) MSR, (d) Soyel's algorithm, (e) proposed algorithm.



(a) (b) (c)

(d) (e)

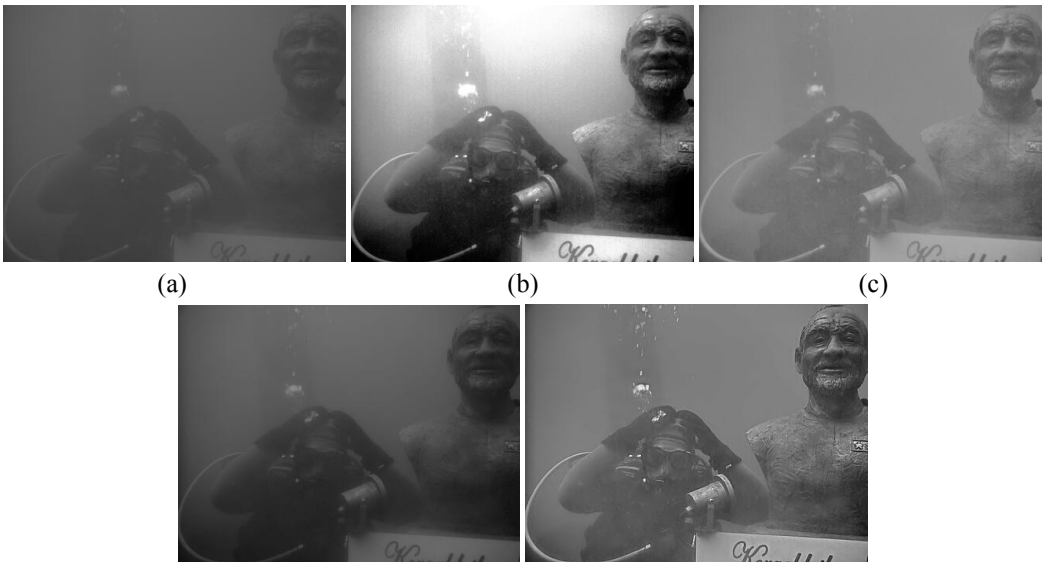
Fig. 7. Enhancement results for image *Testboard*: (a) Original image, (b) HE, (c) MSR, (d) Soyel's algorithm, (e) proposed algorithm.



(a) (b) (c)

(d) (e)

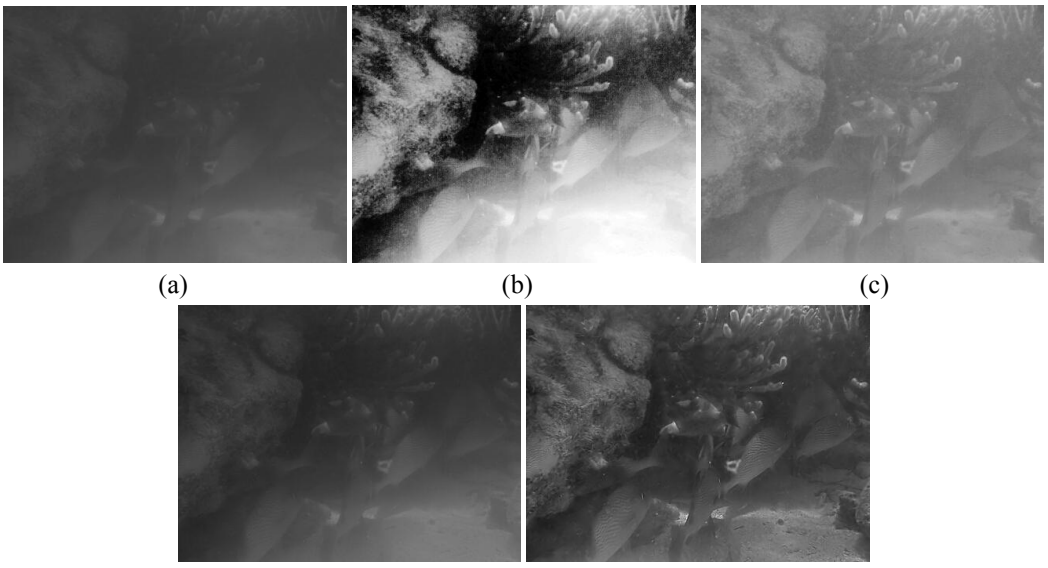
Fig. 8. Enhancement results for image *Archway*: (a) Original image, (b) HE, (c) MSR, (d) Soyel's algorithm, (e) proposed algorithm.



(a) (b) (c)

(d) (e)

Fig. 9. Enhancement results for image *Statue*: (a) Original image, (b) HE, (c) MSR, (d) Soyel's algorithm, (e) proposed algorithm.



(a) (b) (c)

(d) (e)

Fig. 10. Enhancement results for image *Reef*: (a) Original image, (b) HE, (c) MSR, (d) Soyel's algorithm, (e) proposed algorithm.

Table 1. Subjective visual assessment scores as MOSs

Image	Original	HE	MSR	Soyel's algorithm	Proposed
Chair	1.63	2.57	3.02	3.78	4.21
Fish	2.24	3.26	3.95	3.52	4.33
Testboard	1.78	2.63	3.45	3.33	4.54
Archway	1.58	2.11	3.38	3.46	4.29
Statue	1.22	2.03	3.10	2.94	4.56
Reef	1.05	1.67	3.26	3.01	4.19

Table 2. EMEEs of different enhancement algorithms

Image	Original	HE	MSR	Soyel's algorithm	Proposed
Chair	2.27	175.38	324.45	426.50	1722.70
Fish	0.15	251.31	97.12	33.12	792.06
Testboard	0.06	11.71	73.77	12.05	152.63
Archway	0.46	53.34	32.09	1.62	217.37
Statue	0.05	39.31	0.09	0.13	92.95
Reef	0.06	124.05	0.14	0.11	311.49

Table 3. Comparison of running time (s)

Image	HE	MSR	Soyel's algorithm	Proposed
Chair	0.15	0.44	18.95	19.48
Fish	0.16	0.69	23.59	24.10
Testboard	0.09	0.27	8.55	8.89
Archway	0.13	0.38	15.32	15.91
Statue	0.14	0.41	18.22	18.85
Reef	0.14	0.42	18.39	18.91

6. Conclusion

In this paper, a new HVS-based automatic underwater image enhancement algorithm in NSCT domain is proposed. This algorithm combines NSCT transform techniques with the HVS phenomena. To eliminate non-uniform illumination in underwater images, MSR algorithm based on the HVS is integrated into NSCT domain, which can adjust the dynamic range properly. To reduce the noise, threshold denoising is also implemented in NSCT domain. The HVS-based NSCT is then proposed, which utilizes the LM and CM characteristics of the HVS, as well as the advantages of NSCT in capturing the geometric structures of an image. A new nonlinear mapping function is designed to manipulate the HVS-based NSCT contrast coefficients to amplify the edges, and a new nonlinear gain function is used to modify the lowpass subband coefficients to adjust the global dynamic range. Therefore, the proposed algorithm allows for underwater image denoising, edge sharpening and contrast enhancement to be achieved automatically and simultaneously. Experimental results show that the proposed algorithm outperforms the existing enhancement techniques in both MOS and EMEE assessments. However, the proposed algorithm can automatically achieve better results at the cost of a higher computational complexity which is mainly caused by NSCT decomposition.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No.60972101 and No.41306089) and the Natural Science Foundation of Jiangsu Province (No.BK20130240).

References

[1] J. Y. Chiang and Y. C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *IEEE Transactions on image processing*, vol. 21, no. 4, pp. 1756-1769, April, 2012. Article (CrossRef Link)

[2] S. C. Nercessian, K. A. Panetta and S. S. Agaian, "Non-linear Direct Multi-scale Image Enhancement Based on the Luminance and Contrast Masking Characteristics of the Human Visual System," *IEEE Transactions on image processing*, vol. 22, no. 9, pp. 3549-3561, September, 2013. Article (CrossRef Link)

[3] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, Upper Saddle River, NJ, 2006.

[4] T. Celik and T. Tjahjadi, "Automatic image equalization and contrast enhancement using Gaussian mixture modeling," *IEEE Transactions on image processing*, vol. 21, no. 1, pp. 145-156, January, 2012. Article (CrossRef Link)

[5] D. J. Jobson, Z. Rahman and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on image processing*, vol. 6, no. 7, pp. 965-976, July, 1997. Article (CrossRef Link)

[6] D. J. Jobson, Z. Rahman and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Transactions on image processing*, vol. 6, no. 3, pp. 451-462, March, 1997. Article (CrossRef Link)

[7] M. N. Do and M. Vetterli, "Contourlets," *Studies in Computational Mathematics*, vol. 10, pp. 83-105, 2003. Article (CrossRef Link)

[8] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on image processing*, vol. 14, no.12, pp. 2091-2106, December, 2005. Article (CrossRef Link)

[9] J. P. Zhou, A. L. Cunha and M. N. Do, "Nonsubsampled contourlet transform: construction and application in enhancement," in *Proc. of IEEE International Conference on Image Processing*, vol. 1, pp. I-469-72, September 11-14, 2005. Article (CrossRef Link).

[10] A. L. Cunha, J. P. Zhou and M. N. Do, "The nonsubsampled contourlet transform: theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089-3101, October, 2006. Article (CrossRef Link)

[11] S. Li, L. Fang and H. Yin, "Multitemporal image change detection using a detail-enhancing approach with nonsubsampled contourlet transform," *IEEE Geoscience and Remote Sensing Letters*. vol. 9, iss. 5, pp. 836-840, September, 2012. Article (CrossRef Link).

[12] H. J. Li, Z. M. Zhao and X. L. Yu, "Grey theory applied in non-subsampled contourlet transform," *IET Image Processing*, vol. 6, no. 3, pp. 264-272, April, 2012. Article (CrossRef Link)

[13] Q. W. Li, G. Y. Huo, H. Li, G. C. Ma and A. Y. Shi, " Bionic vision-based synthetic aperture radar image edge detection method in non-subsampled contourlet transform domain," *IET Radar Sonar and Navigation*, vol. 6, iss.6, pp.526-535, July, 2012. Article (CrossRef Link)

[14] C. C. Lee, C. Y. Shih, S. K. Lee and W. T. Hong, "Enhancement of blood vessels in retinal imaging using the nonsubsampled contourlet transform," *Multidimensional Systems And Signal Processing*, vol. 23, iss. 4, pp. 423-436, December, 2012. Article (CrossRef Link)

[15] H. Soyel and P. W. McOwan, "Automatic image enhancement using intrinsic geometrical information," *Electronics Letters*, vol. 48, no. 15, pp. 917-919, July, 2012. Article (CrossRef Link)

[16] Z. X. Wang, X. B. Xu, W. Y. Yan, W. Wei, J. H. Li and D. Y. Zhang, "Optimal scheme of retinal image enhancement using curvelet transform and quantum genetic algorithm," *KSII Transactions*

on Internet and Information Systems, vol. 7, no. 11, pp. 2702-2719, November, 2013. Article (CrossRef Link).

[17] F. Yang, Y. Chang and S. Wan, "Gradient-threshold edge detection based on the human visual system," *Opt. Eng.*, vol. 44, no. 2, pp. 020205-1-2, February, 2005. Article (CrossRef Link)

[18] A. Liu, W. Lin, M. Paul, C. Deng and F. Zhang, "Just noticeable difference for images with decomposition model for separating edge and textured regions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 11, pp. 1648-1652, November, 2010. Article (CrossRef Link)

[19] X. K. Yang, W. S. Ling, Z. K. Lu, E. P. Ong and S. S. Yao, "Just noticeable distortion model and its applications in video coding," *Signal Processing: Image Communication*, vol. 20, no. 7, pp. 662-680, August, 2005. Article (CrossRef Link)

[20] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no.6583, pp. 607-609, June, 1996. Article (CrossRef Link)

[21] A. V. Rangan, L. Tao, G. Kovacic and D. Cai, "Multiscale modeling of the primary visual cortex," *IEEE Engineering in Medicine and Biology Magazine*, vol. 28, no. 3, pp. 19-24, May-June, 2009. Article (CrossRef Link)

[22] E. H. Land and J. J. McCann, "Lightness and retinex theory," *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1-11, January, 1971. Article (CrossRef Link)

[23] C. H. Lee, J. L. Shih, C. C. Lien and C. C. Han, "Adaptive Multiscale Retinex for Image Contrast Enhancement," in *Proc. of 2013 International Conference on Signal-Image Technology & Internet-Based Systems*, pp. 43-50, December 2-5, 2013. Article (CrossRef Link).

[24] G. Buchsbaum, "An analytical derivation of visual nonlinearity," *IEEE Transactions on Biomedical Engineering*, vol. BME-27, no. 5, pp. 237-242, May, 1980. Article (CrossRef Link)

[25] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *Journal of the Optical Society of America*, vol. 70, no. 12, pp. 1458-1471, December, 1980. Article (CrossRef Link)

[26] D. L. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425-455, 1994. Article (CrossRef Link)

[27] D. D. Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," *IEEE Transactions on Image Processing*, vol. 15, iss. 6, pp.1610-1620, June, 2006. Article (CrossRef Link)

[28] *Methods for Subjective Determination of Transmission Quality*, ITU-T Recommendation P.800, Geneva, Switzerland, August, 1996.

[29] S. S. Agaian, K. P. Lentz and A. M. Grigoryan, "A new measure of image enhancement," *presented at the IASTED Int. Conf. Signal Processing Communication*, September 19-22, 2000. Article (CrossRef Link).

[30] S. S. Agaian, B. Silver and K. A. Panetta, "Transform coefficient histogram-based image enhancement alogrithms using contrast entropy," *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 741-758, March, 2007. Article (CrossRef Link)

Yan Zhou received the B.S. degree in communication engineering from Xi'an Jiaotong University, China, in 2003 and the M.S. degree in communication and information systems from Xi'an Jiaotong University, China, in 2007. She is a lecturer in Hohai University, China. She is currently pursuing Ph.D. degree in the College of Computer and Information Engineering, Hohai University, China. Her research interests include image processing and pattern recognition.

Qingwu Li received the B.S. degree from Zhengzhou University, China, in 1985, the M.S. degree from Xidian University, China, in 1990 and the Ph.D. degree from Hohai University, China, in 2010. He is a professor and a doctoral supervisor in Hohai University, China. His research interests include digital image processing, information acquisition and intelligent perception, etc.

Guanying Huo received the B.S. degree and the M.S. degree from Xidian University, China, in 2001 and 2004 respectively. He received the Ph.D. degree from Hohai University in 2012. He is an associate professor in Hohai University, China. His research interests focus on sonar image processing.