An objective assessment method for image defogging effects

Fan Guo*, Jin Tang and Zixing Cai

School of Information Science and Engineering,
Central South University,
Changsha Hunan, 410083, China
Email: guofancsu@163.com
Email: tjin@csu.edu.cn
Email: zxcai@csu.edu.cn
*Corresponding author

Abstract: Methods to identify image defogging effects can help evaluate and optimise image defogging algorithms. However, such methods that assess the performance of defogging algorithms or compare them with other algorithms are lacking, whereas existing assessment methods are mostly inadequate. Two new methods to assess defogging effects are therefore proposed in this study. The first method generates synthetic foggy images by using the transmission map of an image degradation model in a full-reference manner. The second method develops an assessment system from the perspective of human visual perception in a no-reference manner. Experimental results from a comparison of defogging algorithms demonstrate the effectiveness and reliability of our proposed methods. Compared with other existing methods, our proposed methods efficiently assess defogging effects from generated synthetic images and human visual perceptions. These methods represent novel approaches to measure defogging effects.

Keywords: image defogging algorithm; defogging effects; image assessment; simulated foggy image; human visual perception.

Reference to this paper should be made as follows: Guo, F., Tang, J. and Cai, Z. (2015) 'An objective assessment method for image defogging effects', *Int. J. Autonomous and Adaptive Communications Systems*, Vol. 8, Nos. 2/3, pp.180–199.

Biographical notes: Fan Guo received her BS degree from the Central South University, Hunan, China, in 2005, MS and PhD degrees in Computer Science from Central South University in 2008 and 2012, respectively. Currently, she is a Lecturer and Postdoctoral Fellow with the School of Information Science and Engineering, Central South University. Her main research interests include image processing, pattern recognition and virtual reality.

Jin Tang received his BS and MS degrees from Peking University, Beijing, China, in 1987 and 1990, respectively, and PhD in Pattern Recognition and Intelligence System from Central South University, Hunan, China, in 2002. He is currently a Professor in the School of Information Science and Engineering, CSU, Changsha. His current research interests are focused on image processing, pattern recognition and computer vision.

Zixing Cai received his diploma degree from the Department of Electrical Engineering, Jiao Tong University, Xi'an, China, in 1962. He has been teaching and doing research at the School of Information Science and

Engineering, Central South University, Changsha, China, since 1962. He has authored/coauthored over 600 papers and 30 books. His current research interests include intelligent systems, artificial intelligence, intelligent computation, and robotics. He received over 30 state, province, and university awards in science, technology, and teaching. One of the most recent prizes is the State Eminent Professor Prize of China.

This paper is a revised and expanded version of a paper entitled 'Objective assessment method for image defogging effect' presented at The 7th International Conference on Frontier of Computer Science and Technology (FCST-12), Suzhou, China, 21–23 November 2012.

1 Introduction

Due to the adverse weather conditions like the presence fog or heavy rain, digital images are easily subjected to a wide variety of disturbance during acquisition, which may reduce visual effect and affect post-processing of the image. Since the importance of the defogging algorithm, much work has been done (Tan, 2008; Tarel and Hautiere, 2009; He et al., 2009). Subjective evaluation, as the main method to assess the efficiency and effectiveness of the enhancement algorithm at present is easily plagued by observer's subjective feeling, which makes the evaluation results unreliable. Specifically, the reasons for the difficulties of defogging effect assessment are:

1 No ideal image as assessment reference image. The assessment for defogging effect is different from the image quality or image restoration assessment, for the sunny day reference image that has completely same scene with foggy image is usually very hard to get.

2 The evaluation criteria for defogging effect should be consistent with human visual perception. However, the human visual perception itself is not a deterministic process. Therefore, it is hard to determine the most important factors that affect human visual decision and design the corresponding assessment indicators.

3 The image quality assessment indexes, such as mean square error (MSE), entropy of information, peak signal to noise ratio (PSNR), are often adopted for defogging effect assessment, which usually get the inconsistent results. Thus, it is very important to research on the objective assessment method of image defogging effect.

Recently, objective image quality evaluation approaches can be classified into three categories: full-reference (Wang et al., 2004; Eskicioglu and Fisher, 1995), reduced-reference (Wang et al., 2004) and no-reference or 'blind' quality assessment approach (Carnec et al., 2008), depending of the requirement for reference image. For the image defogging effect, the objective assessment method can be classified into two categories:

1 Assessing from the image contrast. A typical contrast-based enhancement algorithm is proposed by Hautiere et al. (2006). The method computes the ratio between gradient of the visible edges between the image before and after contrast restoration.

In this way, three indicators of contrast measurement are provided based on the concept of visibility level.

2 Assessing from both image contrast and image colour. Work like Yu et al. (2011b) measures the extent of contrast enhancement from global and local contrast of defogging image. Meanwhile, the hue polar histogram, principal component analysis of RGB image and histogram similarity are used to assess the image colour quality from hue reproduction capacity, colour reproduction capacity and naturalness, respectively. Li et al. (2011) use canny edge detector and bright channel to detect the intensity of effective edge, and also utilises the histogram similarity to denote the colour performance of the defogging image. The notions of detail restoration ratio and colour restoration ratio are proposed in Yao et al. (2009), with which the effect of different defogging algorithms can be also evaluated objectively from the image contrast and colour.

In this paper, we first enumerate the limitations of the existing method, and then propose two assessment methods for image defogging effect by using the full-reference and no-reference method. One is using synthetic foggy image simulated by transmission map, and the other is constructing assessment system from the perspective of human visual perception. The former is a full-reference method which needs images of the same scene with and without fog. However, obtaining such kind of pairs of images is extremely difficult in practice since it requires checking that the illumination conditions are the same into the scene. Thus, a set of synthetic images with and without fog is built up by using virtual reality technique for the defogging effect evaluation. The latter is a no-reference method based on the human visual perception. The method combines three components, contrast, naturalness and colourfulness to yield an overall defogging result measure. Meanwhile, by doing statistic to the series of image sequences from dense fog to over enhancement, the overall variation trend of the three components are analysed to construct a comprehensive evaluation function for assessing the image defogging effect.

Although the two assessment methods have fundamental difference in solutions, since the methods both have pluses and minuses, so which method should be chosen depends on the assessment object and the goal in real situation. Specifically, for the evaluation of the defogging performance of the algorithms on the whole, the synthetic image method should be adopted. While for the comparison with existing defogging algorithms for a given foggy image, the evaluation function method can be used. Compared with other works, our proposed assessment methods have the following advantages:

1 ideal no fog reference image are generated for the assessment method based on synthetic image, which provides a solutions to the problem of the defogging effect assessment

2 the most sensitive factors for subjective assessment are determined from the perspective of human visual perception, and the variation trend of the indexes that reflects these factors are analysed to construct a comprehensive evaluation function, which ensure that the assessment results are consistent with human visual perception.

The remaining of this paper is organised as follows. In Section 2, we summarised the assessment method based on the visibility edges, and discussed its difficulties and limitations. Section 3 is dedicated to the defogging assessment based on our synthetic image generation. Section 4 shows the method based on the human visual perception. Finally, Section 5 presents the experimental results.

2 Defogging effect assessment based on visibility edges

So far, there have been many works to remove fog from image, but only few researches focus on the quantitative measure of defogging effect. In Tarel and Hautiere (2009) and Hautiere et al. (2006), a method for contrast enhancement assessment was proposed, based on the concept of visibility level that is commonly used.

2.1 Theoretical foundation

In the International Commission on Illumination (CIE) Report 19.2, Blackwell introduced the very descriptive term visibility level (*VL*) (CIE, 1987). *VL* is obtained by the ratio of the actual luminance difference the target display to its threshold value, which is defined as:

$$VL = \frac{\Delta L_{actual}}{\Delta L_{threshold}}. \tag{1}$$

where ΔL_{actual} can be estimated by measuring the luminance of the target and its background, and Adrian's empirical target visibility model (Adrian, 1989) can be used to compute the luminance difference threshold $\Delta L_{threshold}$. However, For a complex image which contains several objects on a non-uniform background, to calculate the value of $\Delta L_{threshold}$ is still a challenging task. Fortunately, to solely assess the performance of a contrast restoration method, the approach describe in Hautiere et al. (2006) is much easier. It is proposed to compute the following coefficient *r*:

$$r = \frac{VL_r}{VL_0} = \frac{\Delta L_r}{\Delta I_0}. \tag{2}$$

where VL_r and VL_0 denote the visibility level of the considered object in restored image and the original image respectively. ΔI_r denotes the gradient in the restored image, ΔI_0 the gradient in the original image. Consequently, the computation of r enables to compute the gain of visibility level produced by a defogging algorithm, which measured by the gradient of each pixel belonging to a visible edge in the restored and original image.

2.2 Assessment index

Defogging effect assessment based on visibility edges (Hautiere et al., 2006) obtains contrast map with logarithmic image processing (LIP) model (as shown in Figure 1), and the definition of the meteorological visibility distance proposed by CIE. Three descriptors: the rate *e* of edges newly visible after restoration, the mean $\overline{r}$ over these edges of the ratio of the gradient norms after and before defogging, and the percentage of pixels σ which becomes saturated (black or white) after defogging but were not before are computed to objectively assess defogging effect from different angles.

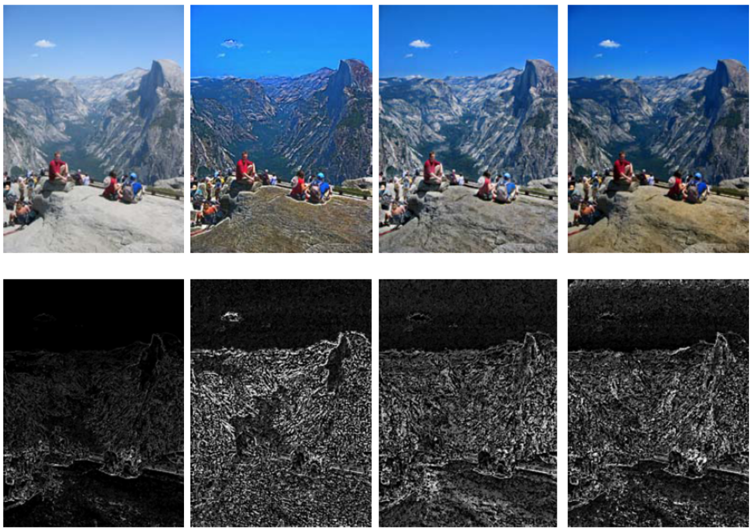
$$e = \frac{n_r - n_0}{n_0}. \tag{3}$$

$$\overline{r} = \exp\left[\frac{1}{n_r}\sum_{P_i \in \wp_r} \log r_i\right]. \tag{4}$$

$$\sigma = \frac{n_s}{dim_x \times dim_y}. \tag{5}$$

where n_0 and n_r denote respectively the cardinal numbers of the set of visible edges in the original image I_0 and in the defogging image I_r. $\wp_r$ is the set of visible edge in I_r. P_i denotes the pixels that belong to the visible edge of image r I , i r is the ratio of the Sobel gradient at P_i in the image I_r and I_0. n_s is the number of pixels which becomes saturated (black or white). dim_x and dim_y denote respectively the width and the height of the image. Figure 1 shows an example of contrast map computation.

Figure 1 The defogging results and their contrast map, (a) Original image (b) Tan's result (c) Tarel's result (d) He's result (see online version for colours)



(a) (b) (c) (d)

Table 1 Three descriptors computed for the compared methods on Figure 1

Descriptor	*Tan's method*	*Tarel's method*	*He's method*
e	–0.08	–0.008	0.06
$\overline{r}$	2.08	2.01	1.42
σ	0.005	0.0	0.002

For each defogging algorithm, the main aim is to increase the contrast without saturating and thus losing some visual information. Hence, the quality of the defogging can be expressed by the three descriptors, and good results are described by high values of e and $\overline{r}$ and low values of σ. Table 1 presents the three descriptors computed for the defogging images in Figure 1. Note that the value of e may be negative. This could be happen when the original image is enhanced over certain extent. Although there are more visibility pixels, these pixels connect together which make the visibility edges become less.

2.3 Limitations

- *The effect definition problem:* The most essential problem with the method in Hautiere et al. (2006) is the definition of defogging effect. In particular, it is not always true that higher values of e or $\bar{r}$, and lower values of σ correspond to a better defogging results. An obvious example would be the over enhancement may be clearly visible but not so objectionable. From Table 1, we deduce that Tan's algorithm is the best because of the highest $\bar{r}$. However, the algorithm probably increased contrast so strong that the image seems have halos near some edges and the colour after defogging also seems unnatural.
- *The inconsistency problem:* When one choose to use one of the descriptors to assess the defogging effect, one is implicitly assuming that the best defogging algorithms should be picked out by any descriptor. In a word, the best algorithm should have the highest e, $\bar{r}$ and lowest σ at the same time. Empirically, however, this is not the case for the state of art algorithms as shown in Table 1. In fact, the best algorithm may be a trade-off among these descriptors. Thus, it remains to be seen how much these descriptors can assess the performance of the current defogging algorithms.

3 Defogging effect assessment based on synthetic image

For objective assessment method, full-reference way is most reliable. However, there is no easy way to have a reference image. Thus, a synthetic image database with and without fog is built up. The database comprises 54 synthetic images of 18 urban road scenes. To each image without fog is associated a foggy images and a transmission map. Fog is added on each image according to image degradation model, and the size of each image is 640 × 480. These scenes are used to test defogging algorithms intensively and in an objective way.

3.1 The generation of synthetic image

The generation of synthetic image makes full use of the depth information to let the distant objects in the scene gradually disappear in the fog. Therefore, the generated fog has very natural visual effect. The image degradation model (Nayar and Narasimhan, 2002) that describes the degradation process and mechanism of the foggy image is the theoretic foundation of the assessment method. The formation of an image degradation model is as following:

$$I(x) - J(x)e^{-\beta d(x)} + A\left(1 - e^{-\beta d(x)}\right) \quad (6)$$

where x is the scene point corresponding to pixel $x = (x, y)$, A is the atmospheric light value, which can be obtained by using the defogging algorithms proposed by Tarel and Hautiere (2009), He et al. (2009), Yu et al. (2011a), etc. $d(x)$ is the distance along the real-word ray corresponding to the pixel x. β is the atmospheric attenuation coefficient due to the light scattering. According to the radiative transport equation proposed by Rossum and Nieuwenhuizn (1999), let $t(x) = e{-}^{kd(x)}$, and $t(x)$ is the transmission map describing the portion of light that is not scattered and reaches the camera.

Considering the distribution of fog usually change with the image depth in real situation, so the parameter λ is introduced to simulate this type of fog. Specifically, a denser medium can be simulated by multiplying the attenuation coefficient β by a factor of λ. This is achieved by applying the following simple power law transformation of the transmission values

$$t'(x) = e^{-\lambda\beta d(x)} = \left(e^{-\beta d(x)}\right)^{\lambda} = t(x)^{\lambda}. \tag{7}$$

Note that the final fog effect obtained by the transformed transmission map $t'(x)$ is mainly depends on the value of parameter λ. Figure 2 shows the simulated foggy images with different λ.

Figure 2 The synthetic foggy images with different by using transmission map, (a) Input image (b) Transmission map (c) $\lambda = 3$ (d) $\lambda = 8$ (see online version for colours)

(a) (b) (c) (d)

3.2 Application example

In order to validate the effectiveness of the defogging effect assessment based on synthetic image, we apply each defogging algorithm on the synthetic fog. Used algorithms are: He's algorithm and Tarel's algorithm. The results on a group of synthetic foggy images are presented in Figure 3. Notice that the increase of the clearness for the farther objects: some object that was barely visible in synthetic foggy image appears clearly in defogging images. A first visual analysis confirms that the enhancement with He's method allows to keep the good properties of the preservation of image detail and object outline, while the results of Tarel's method show a tendency to have halo artifacts. Therefore, He's algorithms performs better compared to Tarel's algorithm for the synthetic foggy images in Figure 3.

Figure 3 Defogging effect comparison for synthetic foggy image, (a) Input image (b) Simulated image (c) Tarel's result (d) He's result (see online version for colours)

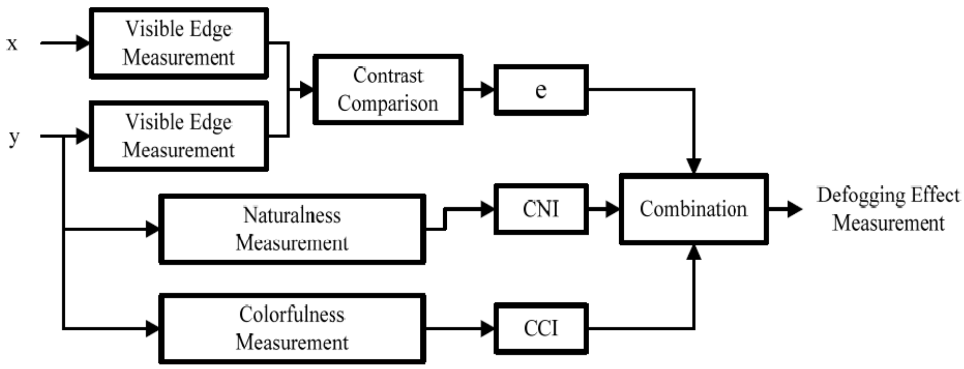
(a) (b) (c) (d)

The quantified comparison consists simply in computing the absolute difference (*AD*) between the image without fog and the image obtained after fog removal. Results, averaged over 18 images are: 47.21 for no defogging operation, 35.07 for Tarel's algorithm, and 31.94 for He's algorithm. One can notice that the *AD* value of He's algorithm is smaller, so the algorithm has better defogging effect since its defogging image is more close to the original image without fog. This confirms our observation on Figure 3.

4 Defogging effect assessment based on human visual perception

The synthetic image method assesses the image defogging effect by using the simulated images of the same scene with and without fog in full-reference way. However, there is no easy way to have the reference image in real situation. Thus, we also try to construct a comprehensive measure system from human visual perception with no need for reference image. Since the three descriptors of the existing assessment method are all based on the contrast map, the quality of the contrast restoration by the defogging algorithms can be well measured. However, human visual system is highly adapted not only to the contrast, but also to the colour quality. Thus, building a comprehensive defogging effect measure system from human visual perception is necessary. In Yendrikhovskij et al. (1998), a model for optimal colour image reproduction of natural images is introduced which are based on the assumption that colour quality of natural images is constrained by perceived naturalness and colourfulness of these images. Therefore, two main factors are also considered here: naturalness and colourfulness. Thus, these three components, contrast, naturalness, colourfulness are combined to yield an overall defogging result measure, which is the contrast-naturalness-colourfulness (CNC) index.

Figure 4 Diagram of the CNC measurement system



4.1 The framework and components of CNC index

The system diagram of the proposed defogging effect assessment system is shown in Figure 4. Suppose *x* is the original foggy image which has unpleasing visual effect, and *y* is the defogging image, then the *CNC* measure can serve as a quantitative measurement of the defogging effect of y. Specifically, the measure system first combines the visibility edges of x and y to compute the contrast measurement index e. Then, computing the image colour naturalness index (*CNI*) and colour colourfulness index (*CCI*). Finally, constructing the comprehensive function with the three indexes, and assessing the image defogging effect by using the function.

The system evaluates the defogging effect from the quality of image contrast and colour, and separates the task of defogging effect measurement into three components: contrast, naturalness and colourfulness. The introduction to the three indexes is as follows:

- *Measurement index for image contrast:* According to the definition of visibility distance (CIE, 1987), the visible edge is the reflection of local contrast, so the contrast can be well measured by the number of visible edges in the images. From (3), one can notice that the value of e may be negative when $n_r < n_0$, which means the number of visible edges after defogging become less. Thus, in order to make the descriptor $e \geq 0$, for the original image x and defogging image y, we define

  $$e(x, y) = \frac{n(y)}{n(x)}. \tag{8}$$

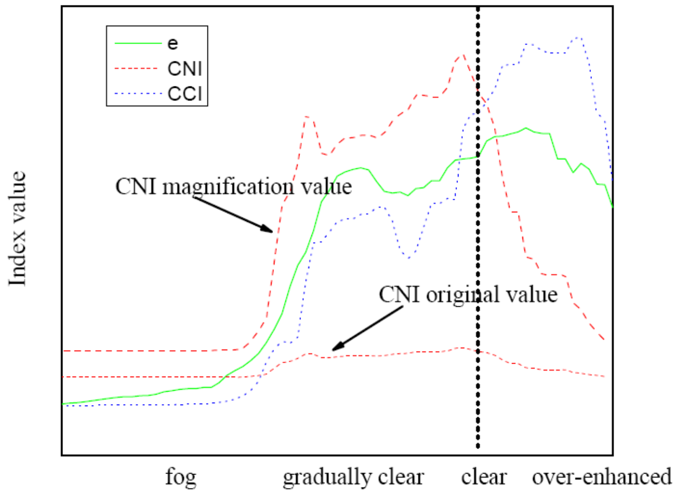
  where $n(x)$ and $n(y)$ denote respectively the cardinal numbers of the set of visible edges in x and y.

- Measurement index for image naturalness: Naturalness is the degree of correspondence between human perception and reality world, which is described by *CNI* (Yendrikhovskij et al., 1998; Huang et al., 2006). We use the index to measure the defogging image y. The bigger value of *CNI* is, the more natural colour image is. To ensure the fairness and objectivity of the evaluation, all the defogging process use the criteria of image segmentation and classification proposed in Huang et al. (2006).

- Measurement index for image colourfulness: Colourfulness presents the colour vividness degree, which is described by *CCI* (Yendrikhovskij et al., 1998; Huang et al., 2006). The index is also used for measuring the defogging image y. When *CCI* is in certain range, the colour of image will be suitable for human. *CCI* is related to image content, and mainly used for measuring the image colourfulness of the same scene in different defogging effect.

In order to analyse the three indexes, we created the simulated foggy image from dense fog, fog, haze, haze-free to over enhancement by gradually adjusting-related algorithm parameters for 50 test images. By doing statistic to the series of image sequences from dense fog to over enhancement, we can obtain the overall variation trend of the three indexes, as illustrated in Figure 5. In Figure 5, green curve represents the variation trend of contrast measurement index e, blue and red curves represent that of *CCI* and *CNI*, respectively. The best defogging effect position is shown in black dotted line. Compared with the change of e and *CCI*, the value change of *CNI* is too small. Thus, the value is enlarged artificially, so the relationship among the indexes can be shown intuitively in the same coordinate system. From Figure 5, we can see that, during the process of gradual clearness, e and *CCI* is ascending at fluctuations. When the image is over-enhanced, the two values are still increasing until the curve increased to a certain degree, it begins to descend rapidly, which means e and *CCI* achieve the best defogging effect before the peak. Here, the fluctuation of e curve is because the visibility pixels that constructed the edges connecting together, which causes the number of visibility edges in the defogging image to change. The fluctuation of *CCI* curve is because of the colour distort during the image enhancement. The curve of *CNI* fluctuant ascending as the increasing of the defogging effect, when the image achieves the best defogging effect, the value begin to decline. *CNI* stands for the naturalness of the image colour. Since the image colour may also be natural when there is a little fog. The curve has some peaks after the image achieves the best defogging effects. Thus, the most natural image is not necessarily has the best clearness effect, but the clear image must have the high CNI value.

Figure 5 The overall variation trend of indexes in CNC system (see online version for colours)



4.2 The construction of comprehensive evaluation function

In practice, one usually requires a single overall defogging effect measure of the entire image. Thus, we combine the three components to yield an overall defogging effect measure

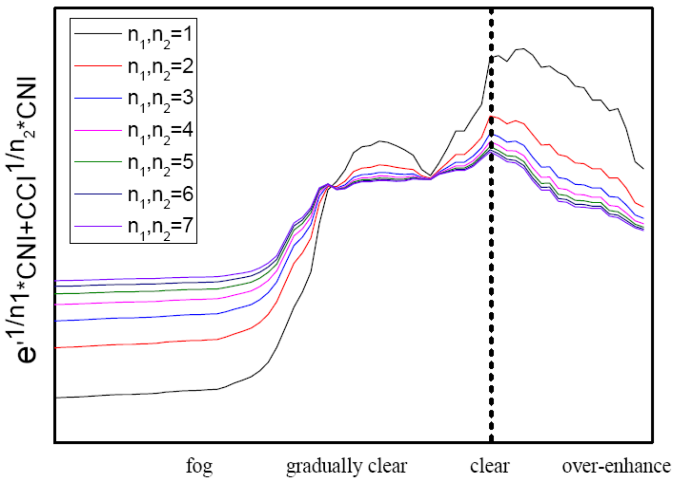
$$CNC(x, y) = f\left(e(x, y), CNI(y), CCI(y)\right). \tag{9}$$

For the overall variation trend of the three indexes shown in Figure 5, we note that, the peak of *CNI* curve stands for the most natural result, but it is not necessarily the best defogging effect. However, the best effect must have good naturalness (high *CNI* value). When the image is over-enhanced, the colour is distorted, and *CNI* is going down rapidly. For *e* and *CCI*, they achieve the best effect before reach their peaks. When the image is over-enhanced, the curves continue ascending. After reach their peaks, these curves just begin to go down. Therefore, if the uptrend of *e* and *CCI* (from their best effect points to their curve's peaks) can be largely counteracted by the downtrend of *CNI*, the peak of *CNC* curve can be more close to the real best effect point. Meanwhile, the value change of *CNI* is small, while that of *e* and *CCI* is relatively big. Thus, the effect of *e* and *CCI* to the *CNC* index needs to be weakened. The *CNC* index between image signals *x* and *y* can be defined

$$CNC(x, y) = e(x, y)^{\frac{1}{n_1}} \cdot CNI(y) + CCI(y)^{\frac{1}{n_2}} \cdot CNI(y). \tag{10}$$

where $n_1 \geq 1$ and $n_2 \geq 1$ are parameters used to weaken the variation trend of e and *CCI*. When n_1 or n_2 is small, the maximum value is achieved when the image is over-enhanced. As the increasing of the two values, the value of $e(x, y)^{1/n_1} \cdot CNI(y)$ and $CCI(y)^{1/n_2} \cdot CNI(y)$ tend to stable, and the peaks of their curve move to the left (the best effect point), see black dotted line. Since the *CNC* index is the sum of the two terms, it has the similar trend, as shown in Figure 6. Therefore, choosing suitable n_1 and n_2 can make *CNC* index stable, and ensure the curve peak is close to the best defogging effect as much as possible. n_1 and n_2 is application-based, we fix $n_1 = n_2 = 5$ in our experiment, the value can make the performance of $e(x, y)^{1/n_1} \cdot CNI(y)$ and $CCI(y)^{1/n_2} \cdot CNI(y)$ keep stable. Thus, the larger the value of *CNC*, the better defogging effect will be.

Figure 6 The overall variation trend curve of CNC by tuning n_1 and n_2 (see online version for colours)



4.3 Application example

We compare some representitive defogging algorithms, such as Tan's, Tarel's and He's method by human perceptual way with our evaluation index. The reason why we choose these algorithms is that they illustrate certain aspects of defogging algorithm. For Tan's method (Tan, 2008), although it can significantly improve image contrast, the results tend to have larger saturation values. We hope these factors can be reflected by our index. The reasons for choosing Tarel's method (Tarel and Hautiere, 2009) and He's method (He et al., 2009) are: The former is one of the fastest defogging algorithms at present, and the latter is recognised as one of the most effective ways to remove fog.

Figures 7 to 9 show some example images by different defogging algorithms. One can notice that, the visual effect of Tan's method is worse than the other two algorithms due to the colour distortion. For Tarel's method and He's method, depending of image, either algorithm could outperform the other. For example, the images produced by Tarel, as shown in Figures 7(c) and 8(c), are more pleasing than He's results. However, for Figure 9, the image generated by He's method may have better defogging effect than Tarel's method. We can see that the traffic sign in this particular image [Figure 9(d)] was better restored with realistic colours. Therefore, from the viewpoint of visual effects, we can pick out the best defogging images: Figure 7(c), Figure 8(c) and Figure 9(d). This can also be testified by our proposed index in Table 2. Good results are described by high value of CNC. From Table 2, we can see that the highest values of CNC are 1.16 (Tarel's method), 1.396 (Tarel's method), and 1.95 (He's method). This confirms our observations on Figures 7 to 9 and indicates that our proposed index delivers better consistency with human visual perception.

Figure 7 Defogging effect comparison for real scene images, (a) Foggy image (b) Tan's result (c) Tarel's result (d) He's result (see online version for colours)

(a) (b) (c) (d)

Figure 8 Defogging effect comparison for real scene images, (a) Foggy image (b) Tan's result (c) Tarel's result (d) He's result (see online version for colours)

(a) (b) (c) (d)

Figure 9 Defogging effect comparison for real scene images, (a) Foggy image (b) Tan's result (c) Tarel's result (d) He's result (see online version for colours)

(a) (b) (c) (d)

Table 2 CNC index computed for the three compared methods

CNC	*Tan's method*	*Tarel's method*	*He's method*
Figure 7	1.13	1.16	1.06
Figure 8	0.94	1.39	1.26
Figure 9	1.49	1.21	1.95

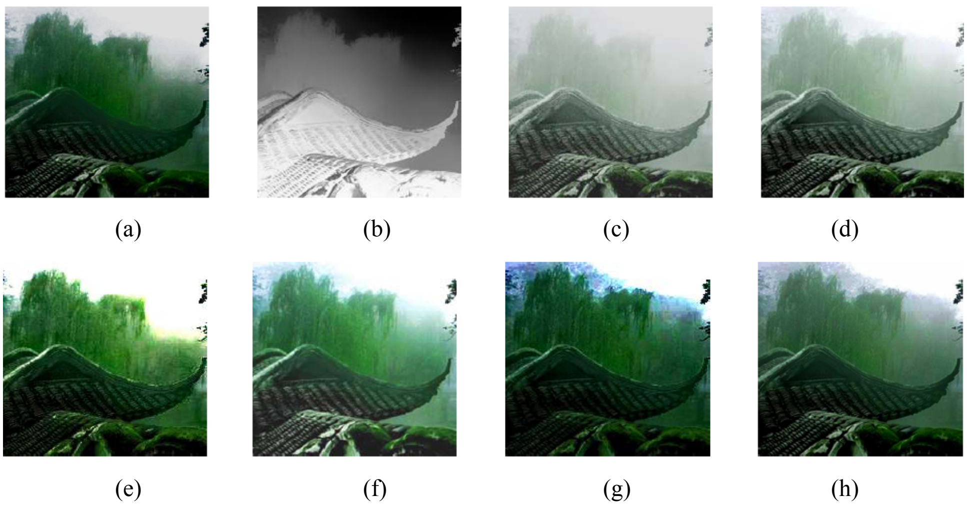
5 Experimental comparison and analysis

In this section, we first use our proposed indexes to assess the defogging effect of different fog removal algorithms. Then, we compare the test results of different assessment methods against a large set of subjective ratings gathered for a database of 80 images to ensure that the qualitative results of defogging algorithms are correctly assessed.

5.1 Test on image database by using our proposed indexes

To validate the effectiveness of the two proposed assessment methods and the simulated fog generation method, experiments are performed with the actual colour images. An illustrative example is shown in Figure 10. Figure 10(a) is the actual no fog image, and its corresponding foggy image is shown in Figure 10(d). The transmission map of the foggy image can be obtained by using the existing defogging algorithms. Figure 10(b) is the transmission map estimated from foggy image [Figure 10(d)] by using Yu's algorithm (Yu et al., 2011a). Using (7), we can get the transformed transmission map. By taking the transformed map, atmospheric light and original no fog image into the expression of image degradation model, the simulated foggy image can be obtained, as shown in Figure 10(c). Figures 10(e) and 10(f) are the defogging results of Tarel's and He's method for simulated foggy image [Figure 10(c)]. Figures 10(g) and 10(h) are the defogging results for actual foggy image [Figure 10(d)].

Figure 10 Defogging effect comparison for synthetic foggy images and real scene images, (a) No fog image (b) Transmission map (c) Simulated foggy image (d) Actual foggy image (e) Tarel's result using (c) (f) He's result using (c) (g) Tarel's result using (d) (h) He's result using (d) (see online version for colours)



(a) (b) (c) (d)

(e) (f) (g) (h)

One can notice that the result of He's method seems more natural than Tarel's result. The statistics of *AD* and *CNC* in Table 3 also confirm this conclusion. However, the contradictory conclusion is drawn by using the visibility edge method. From Table 4, we can see that for simulated foggy image, He's result is better than Tarel's. While for actual foggy image, Tarel's result is better. The reason is because the existing method only assesses the defogging effect from the image contrast without considering the image colour and other important evaluation factors.

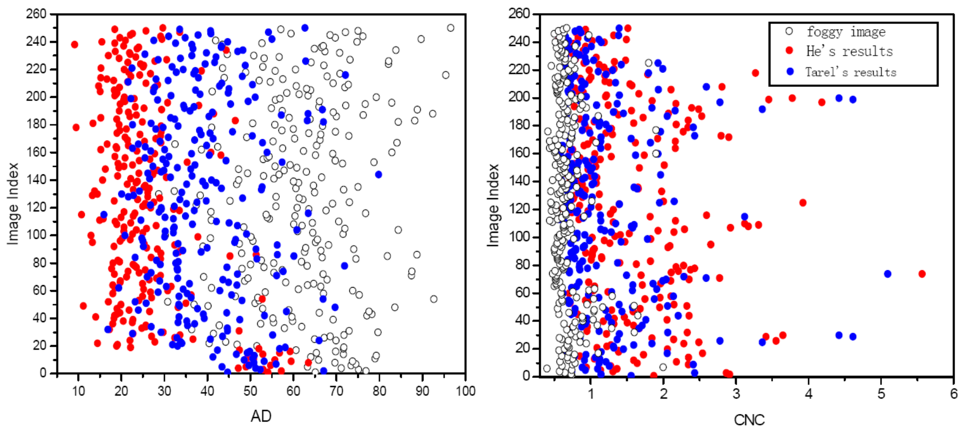
Table 3 Index computed for the two compared algorithms by using the proposed assessment methods

Index	*Nothing (simulated fog)*	*Tarel's method (simulated fog)*	*He's method (simulated fog)*	*Tarel's method (actual fog)*	*He's method (actual fog)*
AD	73.01	45.75	22.94	20.08	15.92
CNC	—	2.52	2.74	2.59	2.78

Table 4 Results of contrast enhancement assessment method on Figure 10

Index	*Tarel's method*		*He's method*	
	Simulated defogging image	*Actual defogging image*	*Simulated defogging image*	*Actual defogging image*
e	0.9256	1.1460	0.4398	0.5656
$\bar{r}$	1.5789	1.0320	1.8873	0.7885
σ (%)	0.9470	1.3990	0	1.6944

Figure 11 Index statistical results for test images (see online version for colours)



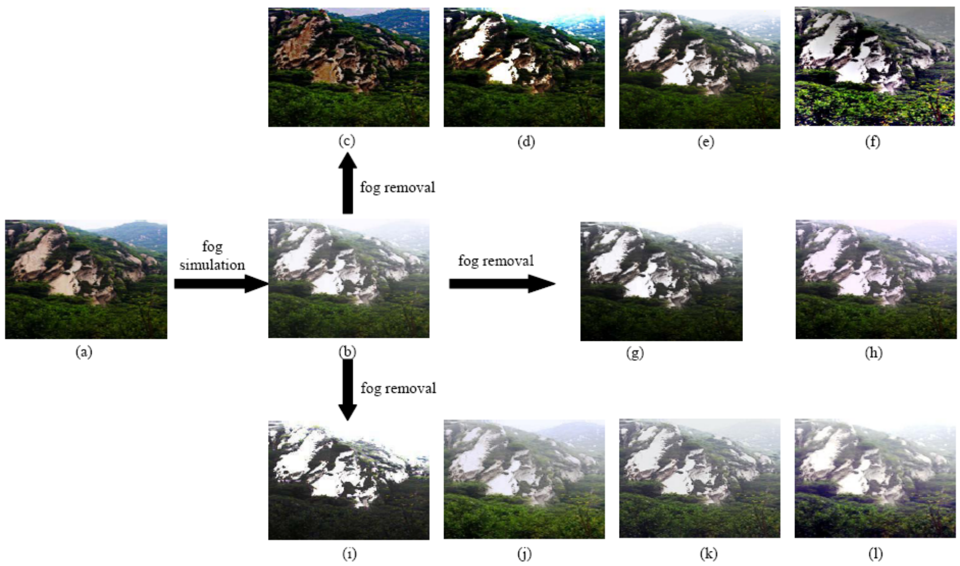
(a) (b)

We also test our assessment index on more sample natural colour images (500 test images), which consisted of foggy image from internet database and real scene captured by canon S80 with different scene, weather and fog density, as shown in Figure 11. Where Figure 11(a) is the statistical result of *AD* for 250 images, and Figure 11(b) is the *CNC* result for other 250 images. In these figures, circle '○' stands for original foggy image, red circle '●' stands for defogging results of He's method and blue circle '●' stands for Tarel's results. The horizontal axes are the assessment index values and vertical axes are the image number index. From Figure 11(a), we can see that the *AD* of He's method is smaller than that of Tarel's method. Since the smaller value *AD* is, the better the defogging effect will be. Thus, the conclusion that He's method has better clearness effect can be drawn. For Figure 11(b), it is clear that the CNC of original foggy image is clustered between 0 and 2, while the *CNC* of Tarel's and He's results are distributed between 0.5 and 5.5, 0.5 and 6, respectively. Since the higher value *CNC* is, the better the defogging effect will be. Thus, we can deduce that the visual effect of original image can be effectively improved by using the two defogging algorithms, and one can notice that the overall value of *CNC* of He's method is a little greater than that of Tarel's method. This indicates that, on the whole, He's defogging effect is slightly better for this image test database, which is consistent with the assessment results of *AD* and human visual perception.

5.2 Subjective vs. objective test for different assessment methods

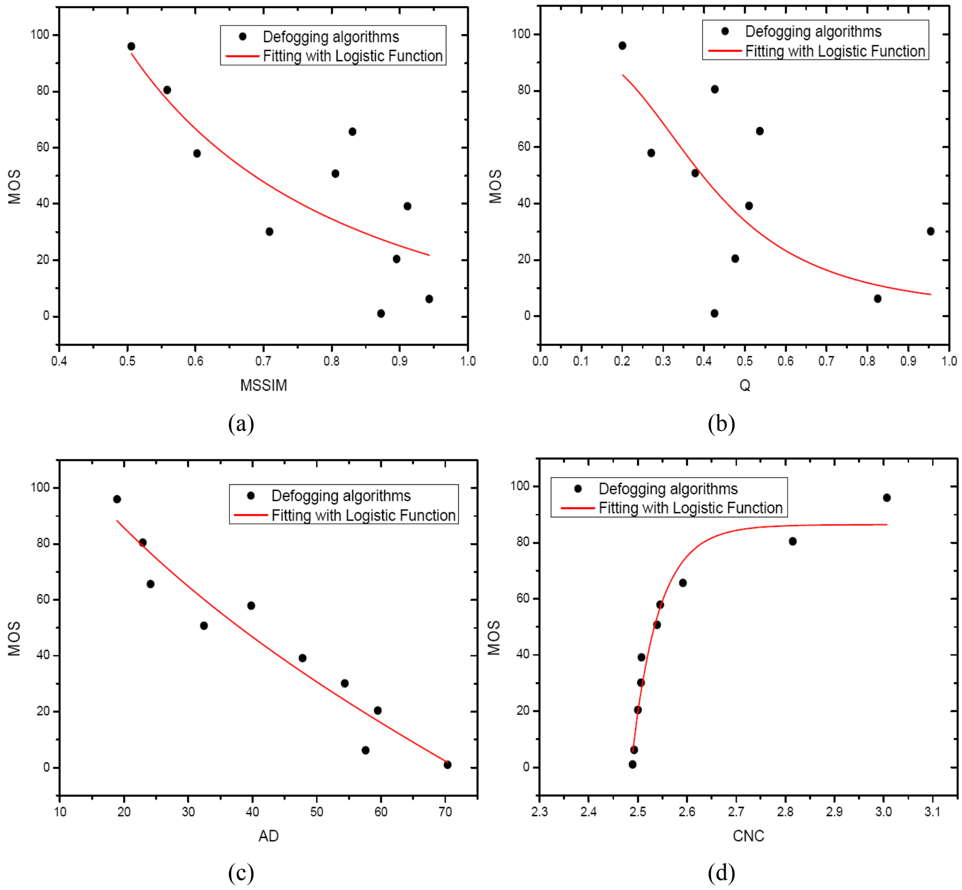
We compare the defogging performance of different assessment methods on a simulated foggy image, as shown in Figure 12. The image were enhanced or restored by using different defogging algorithms, such as Tarel's work (Tarel and Hautiere, 2009), Yu's work (Yu et al., 2011a), He's work (He et al., 2009), Retinex algorithm (Jobson et al., 1997), Fog veil algorithm (Guo et al., 2011), Xu's work (Xu and Xiao, 2009), Fusion algorithm (Ancuti et al., 2010), histogram equalisation (Xu et al., 2009), homomorphic filter (Seow and Asari, 2006) and wavelet transform (John and Wilscy, 2009).

Figure 12 Simulated foggy image and its corresponding defogging images obtained by using different fog removal algorithms, (a) Original no-fog image (b) Simulated foggy image (c) Tarel's result (Tarel and Hautiere, 2009) (d) Yu's result (Yu et al., 2011a) (e) He's result (He et al., 2009) (f) Retinex algorithm (Jobson et al., 1997) (g) Fog veil algorithm (Guo et al., 2011) (h) Xu's result (Xu and Xiao, 2009) (i) Fusion algorithm (Ancuti et al., 2010) (j) Histogram equalisation (Xu et al., 2009) (k) Homomorphic filter (Seow and Asari, 2006) (l) Wavelet transform (John and Wilscy, 2009) (see online version for colours)



Subjects viewed the defogging images [Figure 12(c) to Figure 12(l)] from comfortable seating distances (this distance was only moderately controlled, to allow the data to reflect natural viewing conditions), and were asked to provide their perception of quality on a continuous linear scale that was divided into five equal regions marked with adjectives ‘bad’, ‘poor’, ‘fair’, ‘good’, and ‘excellent’. The test image was views by 18 subjects. The subjects were mostly male college students. Raw scores for each subject were normalised by the mean and variance of scores for that subject [i.e., raw values were converted to Z_scores (van Dijk et al., 1995)] and then the entire dataset was rescaled to fill the range from 1 to 100.

Figure 13 Scatter plots of subjective mean opinion score (MOS) versus assessment indexes for the test image, (a) MSSIM vs. MOS (Wang et al., 2004) (b) Q vs. MOS (Li et al., 2011) (c) AD vs. MOS (d) CNC vs. MOS (see online version for colours)



Notes: Each sample point represents one defogging result. The defogging algorithms are listed as follows: Tarel’s work (Tarel and Hautiere, 2009), Yu’s work (Yu et al., 2011a), He’s work (He et al., 2009), Retinex algorithm (Jobson et al., 1997), fog veil algorithm (Guo et al., 2011), Xu’s work (Xu and Xiao, 2009), fusion algorithm (Ancuti et al., 2010), histogram equalisation (Xu et al., 2009), homomorphic filter (Seow and Asari, 2006) and wavelet transform (John and Wilscy, 2009).

The defogging effect assessment approaches used for comparison include the MSSIM (Wang et al., 2004), Q (Li et al., 2011), our proposed AD and CNC. The scatter plot of MOS versus prediction for each assessment method is shown in Figure 13. As can be seen from the figure, the MSSIM index and the Q index do not deliver satisfactory results in Figure 13. The scatter plots of the two indexes have a dispersive distribution, and do not show any obvious law of variation. Thus, their assessment scores do not really match the view by human. We think the major reason is that the MSSIM, as a image quality assessment index, is not suitable for assessing defogging effect, and the denominator of the Q index is close to zero at nearly haze-free regions, which makes the two indexes unstable. Our proposed AD and CNC perform quite well in this test. By assessing from the no-fog reference images or the perspective of human visual perception, the proposed indexes completely avoid the above problems and the distribution of their scatter plots shows a good monotonicity and function fitting results. Therefore, the scatter slot demonstrates that they supply remarkably good prediction of the subjective scores.

We also compare the different defogging methods on an image database composed of the images from internet and natural images captured by our camera by using the four indexes. 100 simulated foggy image and their corresponding defogging images obtained by using the above ten fog removal algorithms were tested to ensure that the qualitative results of defogging algorithms are correctly assessed and consistent with human visual perception. Experimental results on the images demonstrate that similar distribution trends are obtained, which means our proposed indexes (AD and CNC) perform much better than the previous indexes (MSSIM and Q).

In order to provide quantitative measures on the performance of the objective assessment methods for the 100 test images and their defogging images, we follow the performance evaluation procedures employed in the Video Quality Expert Group (VQEG, 2000) phase I FR-TV test, were four evaluation metrics were used. First, logistic functions are used in a fitting procedure to provide a nonlinear mapping between objective/subjective scores. The fitted curves of the scatter plots for these images have the similar distribution trends as the curves shown in Figure 13.

In VQEG (2000), metric 1 is the correlation coefficient between objective/subjective scores after variance-weighted regression analysis. Metric 2 is the correlation coefficient between objective/subjective scores after nonlinear regression analysis. These two metrics combines, provide an evaluation of prediction accuracy. The third metric is the Spearman rank-order correlation coefficient between the objective/subjective scores. It is considered as a measure of prediction monotonicity. Finally, metric 4 is the outlier ratio (percentage of the number of predictions outside the range of ±2 times of the standard deviations) of the predictions after the nonlinear mapping, which is a measure of the prediction consistency. More details on these metrics can be found in VQEG (2000). In addition to these, we also calculated the mean absolute prediction error (MAE), and root mean square prediction error (RMS) after nonlinear regression, and weighted mean absolute prediction error (WMAE) and weighted root mean square prediction error (WRMS) after variance-weighted regression. The evaluation results for all the assessment methods being compared are given in Table 5. For every one of these criteria, CNC and AD perform better than all of the other methods being compared.

Table 5 Performance comparison of defogging effect assessment methods

Index	Nonlinear regression				Variance-weighted regression				Rank-order
	CC	*MAE*	*RMS*	*OR*	*CC*	*WMAE*	*WRMS*	*OR*	*SROCC*
MSSIM	0.77	5.98	7.68	0.18	0.77	5.73	7.31	0.18	0.77
Q	0.46	8.15	9.98	0.21	0.46	8.09	9.86	0.21	0.44
AD	0.97	3.18	5.13	0.05	0.97	3.09	5.07	0.05	0.98
CNC	0.98	2.89	4.51	0.03	0.98	2.65	4.39	0.03	0.99

Notes: CC: correlation coefficient; MAE: mean absolute error, RMS: root mean squared error; OR: outlier ratio; WMAE: weighted mean absolute error; WRMS: weighted root mean squared error; SROCC: Spearman rank-order correlation coefficient.

6 Conclusions

Image defogging is an important issue in image processing area. Although the methodology of assessing the algorithms or rate them is rare in the literature, the research has significant meaning for improving the quality of the defogging image. In this paper, we have summarised the existing method to image defogging effect assessment based on visibility edges, and have enumerated its limitations. We have proposed two methods based on synthetic foggy image or human visual perception for assessing the defogging effect, which show some new ways of work for defogging measurement. Results on a variety of test images demonstrate the effectiveness and reliability of the proposed methods. Our future work will be focused on the methodology of automatic parameter tuning for the defogging algorithm, so the static open-loop parameter estimation issue can be transformed into the dynamical close-loop parameter tuning issue.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China under Grant 90820302 and 60805027, the Postdoctoral Science Foundation of Central South University (No. 126648). The authors would like to thank Erik Matlin and Kaelan Yee for providing He's source code, Dr. Jean-Philippe Tarel and Dr. Nicolas Hautiere for providing the matlab code of their approach, and Dr. Tan for providing his defogging image. The authors would also like to thank the reviewers and the editor for their valuable comments.

References

Adrian, W. (1989) 'Visibility of targets: model for calculation', *Lighting Research & Technology*, Vol. 21, No. 4, pp.181–188.

Ancuti, C.O., Ancuti, C. and Bekaert, P. (2010) 'Effective single image dehazing by fusion', *Proceedings of the IEEE 17th International Conference on Image Processing*, IEEE, Hong Kong, pp.3541–3544.

Carnec, M., Callet, P.L. and Barba, D. (2008) 'Objective quality assessment of color images based on a generic perceptual reference', *Image Communication*, Vol. 23, No. 4, pp.239–256.

CIE (1987) *International Lighting Vocabulary*, No. 17.4, Paris, France.

Eskicioglu, A.M. and Fisher, P.S. (1995) 'Image quality measures and their performance', *IEEE Trans. Communication*, Vol. 43, No. 12, pp.2959–2965.

Guo, F., Cai, Z.X. and Xie, B. (2011) 'Video defogging algorithm based onfog veil', *ACTA Electronica Sinica*, Vol. 39, No. 9, pp.2019–2025.

Hautiere, N., Tarel, J.P., Aubert, D. and Dumont, E. (2006) 'Blind contrast enhancement assessment by gradient ratioing at visible edges', *Image Analysis & Stereology Journal*, Vol. 27, No. 2, pp.87–95.

He, K.M., J. Sun, and Tang, X.O. (2009) 'Single image haze removal using dark channel prior', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Miami, USA, IEEE, pp.1956–1963.

Huang, K.Q., Wang, Q. and Wu, Z.Y. (2006) 'Natural color image enhancement and evaluation algorithm based on human visual system', *Computer Vision and Image Understanding*, Vol. 103, No. 1, pp.52–63.

Jobson, D.J., Rahman, Z.U. and Woodell, G.A. (1997) 'A muhiseale Retinex for bridging the gap between color images and the human observation of scenes', *IEEE Trans on Image Processing*, Vol. 6, No. 7, pp.969–976.

John, J. and Wilscy, M. (2009) 'Enhancement of weather degraded color images and video sequences using wavelet fusion', *Lecture Notes in Electrical Engineering*, Vol. 39, No. 9, pp.99–109.

Li, D.P., Yu, J. and Xiao, C.B. (2011) 'No-reference quality assessment method for defogged images', *Journal of Image and Graphics*, Vol. 16, No. 9, pp.1753–1757.

Nayar, S.K. and Narasimhan, S.G. (2002) 'Vision in bad weather', *Proceedings of the IEEE International Conference on Computer Vision*, IEEE, Kerkyra, Greece, pp.820–827.

Rossum, Z. and Nieuwenhuizn, T. (1999) 'Multiple scattering of classical waves: microscopy, mesoscopy and diffusion', *Reviews of Modern Physics*, Vol. 71, No. 1, pp.313–371.

Seow, M.J. and Asari, V.K. (2006) 'Ratio rule and himomorphic filter for enhancement of digital color image', *Neurocomputing*, Vol. 69, No. 7, pp.954–958.

Tan, R.T. (2008) 'Visibility in bad weather from a single image', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Anchorgae, USA, pp.1–8.

Tarel, J.P. and Hautiere, N. (2009) 'Fast visibility restoration from a single color or gray level image', *Proceedings of the 12th IEEE International Conference on Computer Vision*, Kyoto, Japan, IEEE, pp.2201–2208.

van Dijk, A.M., Martens, J.B. and Watson, A.B. (1995) 'Quality assessment of coded images using numerical category scaling', *Proc. SPIE*, Vol. 2451.

VQEG (2000) *Final Report from the Video Quality Expert Group on the Validation of Objective Models of Video Quality Assessment*, [online] ftp://vqeg.its.bldrdoc.gov/Documents/Meetings/Hillsboro_VQEG_Mar_03/VQEGIIDraftReportv2a.pdf (accessed 15 January 2012).

Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. (2004) 'Image quality assessment: from error visibility to structural similarity', *IEEE Trans. Image Processing*, Vol. 13, No. 4, pp.600–612.

Xu, D.B. and Xiao, C.B. (2009) 'Color-preserving defog method for foggy or haze scenes', *Proceedings of the 4th International Conference on Computer Vision Theory and Applications*, IEEE, Algarve, Portugal, pp.69–73.

Xu, Z.Y., Liu, X.M. and Chen, X.N. (2009) 'Fog removal from video sequences using contrast limited adaptive hisotogram equalization', *Proceedings of International Conference on Computational Intelligence and Software Engineering*, IEEE, Wuhan, China, pp.1–4.

Yao, B., Huang, L. and Liu, C.P. (2009) 'Research on an objective method to compare the quality of defogged images', *Proceedings of Chinese Conference on Pattern Recognition*, IEEE, Nanjing, China, pp.1–5.

Yendrikhovskij, S., Blommaert, F. and de Ridder, H. (1998) 'Perceptual optimal color reproduction', *Proceedings of SPIE*, pp.274–281.

Yu, J., Li, D.P. and Liao, Q.M. (2011a) 'Physics-based fast single image fog removal', *ACTA Automatics Sinica*, Vol. 37, No. 2, pp.143–149.

Yu, J., Xu, D.B. and Liao, Q.M. (2011b) 'Image defogging: a survey', *Journal of Image and Graphics*, Vol. 16, No. 9, pp.1561–1576.