REFERENCELESS PERCEPTUAL IMAGE DEFOGGING

Lark Kwon Choi[1], *Jaehee You*[2], *and Alan C. Bovik*[1]

[1]Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, USA
[2]Department of Electronic and Electrical Engineering, Hongik University, Seoul, Korea
larkkwonchoi@utexas.edu, jaeheeu@hongik.ac.kr, bovik@ece.utexas.edu

***Abstract*—We propose a referenceless perceptual defog and visibility enhancement model based on multiscale "fog aware" statistical features. Our model operates on a single foggy image and uses a set of "fog aware" weight maps to improve the visibility of foggy regions. The proposed defog and visibility enhancer makes use of statistical regularities observed in foggy and fog-free images to extract the most visible information from three processed image results: one white balanced and two contrast enhanced images. Perceptual fog density, fog aware luminance, contrast, saturation, chrominance, and saliency weight maps smoothly blend these via a Laplacian pyramid. Evaluation on a variety of foggy images shows that the proposed model achieves better results for darker, denser foggy images as well as on standard defog test images.**

Keywords-defog, visibility enhancement, fog aware

I. INTRODUCTION

The perception of outdoor natural scenes is important for successfully conducting visual activities such as object detection, recognition, and navigation. In bad weather, the absorption or scattering of light by atmospheric particles such as fog, haze, and mist can seriously degrade visibility [1]. As a result, objects in images captured under such conditions suffer from low contrast, faint color, and shifted luminance. Since degraded visibility can cause operator misjudgments in vehicles guided by camera images and can induce erroneous sensing in surveillance systems, automatic methods for visibility enhancement of foggy images have been intensively studied [1-9].

The earliest approaches used multiple images of the same scene under different weather conditions to compute a depth map [1] or different degrees of polarization by rotating a polarizing filter attached to a camera [2]. However, acquiring enough images is time-consuming, and it is difficult to find the maximum and minimum degree of polarization during rapid scene changes.

The second approach is to combine a single image with additional depth information obtained either by user input or a 3D geometric model [3]. While this approach avoids the multiple image requirement, it is still difficult to apply in practice because user interaction is not an automatic process, and it is difficult to generate accurate 3D geometric models that can capture dynamic real-world structure.

The third approach is to use only a single image. Tan [4] predicted scene albedo by maximizing the local contrast while assuming a smooth layer of airlight, but the results were overly saturated by halo effects. Fattal [5] improved visibility by supposing that transmission and surface shading are statistically uncorrelated. However, this method requires substantial color and luminance variation to occur in the foggy scene. He *et al.* [6] made the important contribution of the *dark channel prior*. It attains successful results by refining the initial transmission map using a soft matting technique; however, soft matting is computationally quite expensive although it can be sped up using a guided filter [7]. Tarel and Hautière [8] built a fast solution using an edge preserving *median of median filter*, but the extracted depth-map must be smooth except along edges with large jumps.

Recently, Ancuti *et al.* [9] used multi-scale fusion [10-11] for single image dehazing. Image fusion is a way to blend several images into a single one by retaining only the most useful features [10]. Dehazing by fusion has several advantages: it can reduce patch-based artifacts by singe pixel operations, and it is fast since it does not estimate a transmission map. Still, the design of the preprocess images and weight maps from only a single foggy image without other references such as a corresponding fog-free image or geographical information remains difficult.

Ancuti *et al.* derived a second preprocessed image by subtracting the average luminance of a single foggy image, then magnified the difference. This method captures rough haze regions and recovers visibility, but the performance is decreased when the foggy image is dark because the severe dark aspects of the preprocessed image begin to dominate as shown in Figure 2(c). Although saturation, chrominance, and saliency weight maps can help mitigate the degradation, the visibility is not enhanced much as can be seen in Figure 4.

We propose a referenceless multiscale perceptual defog and visibility enhancement model. "Referenceless" means that the proposed model does not require multiple images, side information, and content dependent assumptions such as smoothness of airlight layers, color, depth, even computation of a depth dependent transmission map. While Ancuti *et al.* used only the average luminance of the entire image, we use fog aware statistical features [12] to capture accurate fog-free and foggy regions. The proposed weight maps also include perceptually relevant fog density, fog aware luminance, and contrast attributes. Results on a wide range of foggy images show that the proposed model achieves better results on dark, dense foggy images as well as on standard defog test images.

The rest of this paper is organized as follows. Section 2 reviews the optical model of foggy images. The referenceless defog and visibility enhancement model is described in Section 3. Section 4 studies the performance of the method. We conclude the paper with future work in Section 5.

II. Optical Model of Foggy Image Formation

When light from the sun passes through a scattering fog atmosphere, light reflected from objects is attenuated along the path to the camera and is also scattered in other directions. Using Koschmieder's atmospheric scattering model [13], a foggy image I can be decomposed into two components, direct attenuation and airlight:

$$I(x) = J(x)t(x) + A[1 - t(x)]. \quad (1)$$

Here, $J(x)$ is the scene radiance or fog-free image to be reconstructed, $t(x)$ is the medium transmission at each pixel x, and A is the skylight. This model assumes a linear correlation between the reflected light and the distance between the object and the camera. The first term is direct attenuation representing how the scene radiance is attenuated in the medium. The second term, known as airlight, arises from previously scattered light and causes a shift in scene color. The transmission $t(x)$ can be expressed $t(x) = \exp[-\beta d(x)]$, where β is the medium attenuation coefficient, and $d(x)$ is the distance between the scene and the observer.

III. Perceptual Defogging

Our model executes a defogging process without estimating a transmission map. First, we preprocess a single foggy image in three different ways. Next, six fog aware weight maps are produced from each preprocessed result, then the weight maps are normalized. Finally, the defogged image is obtained via multi-scale fusion, using a Laplacian pyramid. Figure 1 shows a block diagram of the model, and each stage of processing is detailed in the following.

A. Preprocessing

The first preprocessed image, I_1, is white balanced to adjust the natural rendition of the output by eliminating chromatic casts caused by atmospheric color. The shades-of-gray color constancy technique [14] is used, similar to [9], because it is fast and robust when applied to foggy images.

The second and the third preprocessed images are contrast enhanced images. Ancuti *et al.* derived a contrast enhanced image by subtracting the average luminance value, $\bar{I}$, of the image I from the foggy image I, then applying a multiplicative gain. Thus $I_2 = \gamma(I - \bar{I})$, where $\gamma = 2.5$ [9]. Although $\bar{I}$ is a good estimate of image brightness, problems can arise in very dark image regions or in denser foggy images. To overcome this limitation, we also create another type of preprocessed image using a model of statistical regularity observed in natural fog-free images,

$$I_3 = \gamma[I - \mu(I_{fogfree})], \quad (2)$$

where $\mu(I_{fogfree})$ is the average luminance of the fog-free regions only of I. $I_{fogfree}$ indicates where each feature f_i of I takes larger values than $\bar{f}_i$, where $\bar{f}_i = 1/K \times \sum_{k=1}^{K} f_i(k)$ and where $f_i(k)$ is the i^{th} feature of the k^{th} corpus image, and $K = 160$. $f_1 \sim f_8$ include the sharpness, the variance of the mean subtracted contrast normalized (MSCN) image [16], the contrast, the image entropy, the colorfulness, and the color saturation. For f_9, pixel-wise dark channel prior, the regions of interest are instead where the feature f_9 of I takes smaller values than $\bar{f}_9$. When there is no fog-free region in I, the least foggy regions are used. These regions are defined as having larger feature values than the 95% of one or more of $\bar{f}_1 \sim \bar{f}_8$, $(1 - \bar{f}_9)$.When the 95% threshold fails to find the least foggy regions (e.g., on an extremely foggy scene), the percentage is reduced by 5% iteratively until such regions occur. Figure 2 shows a foggy image and the corresponding preprocessed images.

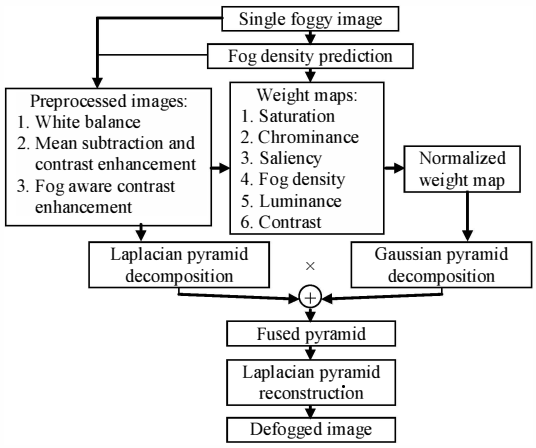


Figure 1. Block diagram of the proposed perceptual defogging model.

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

Figure 2. Original foggy image and preprocessing results. (a) foggy image, I (b) white balanced image, I_1, (c) contrast enhanced image after mean subtraction, I_2 [9], and (d) fog aware contrast enhanced image, I_3.

B. Weight maps

Weight maps selectively filter the most visible regions of the preprocessed images. Ancuti *et al.*[9] used saturation, chrominance, and saliency weight maps. We summarize these and propose a set of additional fog aware weight maps.

The ***saturation weight map***, W_{sat}, measures the visibility of each pixel by estimating the loss of colorfulness. It takes higher values at saturated pixels assumed to be part of haze-free regions. The ***chrominance weight map***, W_{chr}, controls saturation gain with the distance between local saturation S and the maximum ($S_{max} = 1$) in *HSI* color space. The ***saliency weight map***, W_{sal}, emphasizes areas by enhancing the global and local contrast. The maps are computed as follow [9]:

$$W_{sat}^k = \sqrt{1/3[(R^k - L^k)^2 + (G^k - L^k)^2 + (B^k - L^k)^2]},$$
$$W_{chr}^k = \exp\left(-\frac{(S^k - S_{\max}^k)^2}{2\sigma^2}\right), \quad (3)$$
$$W_{sal}^k = \| I_k^{w_{hc}} - I_k^{\mu} \|,$$

where k is an index on the preprocessed images. R^k, G^k, B^k, and L^k are the red, green, blue color channels and the average luminance of I_k. The standard deviation, $\sigma = 0.3$ [9]. $I_k^{w_{hc}}$ is a Gaussian smoothed version of I_k, I_k^{μ} is the mean pixel value of I_k in *Lab* color space, and $\| \|$ is the L_2 norm [15].

The ***fog density weight map*** plays an integral role of guiding the other weight maps to accurately select and filter fog-free and foggy regions. A perceptual fog density map on I is first predicted using a Mahalanobis-like distance measure on overlapped 8 × 8 patches [12]. Next, a guided filter [7] is applied to reduce noise, and the range of the denoised fog density map is scaled to [0 1]. As shown in Figure 2, since I_2 contains the most visible information regarding denser foggy regions of I, we assign the denoised and scaled fog density map, D_d, to the fog density weight map of I_2 as follows:

$$W^2_{fog} = D_d,\ W^1_{fog} = 1 - D_d,\ W^3_{fog} = W^1_{fog} \times W^2_{fog}, \tag{4}$$

where W^3_{fog} is scaled to [0 1].

The ***fog aware luminance weight map*** represents how close the luminance of the preprocessed images is to the luminance of the fog-free (or least foggy) regions of I. Since contrast enhancement methods often cause severe shifts in the luminance profiles of the processed images, yielding excessively dark patches or a faded appearance in some areas, the fog aware luminance weight map seeks to alleviate these degradations by allocating a high value to luminances closer to $\mu(I_{fogfree})$. The map is achieved using a Gaussian curve at each RGB color channel, then they are multiplied,

$$W^k_{lum} = W^k_{lum\_R} \times W^k_{lum\_G} \times W^k_{lum\_B},$$
$$W^k_{lum\_i} = \exp\left(-\frac{[I^i_k - \mu(I^i_{fogfree})]^2}{2\sigma^2} \right), \tag{5}$$

where I^i_k is the color channel of I_k, and $\mu(I^i_{fogfree})$ is the average luminance of $I^i_{fogfree}$ at $i \in \{R, G, B\}$. $\sigma = 0.2$ [11].

The ***contrast weight map*** indicates the sharpness of the preprocessed images by assigning higher weights at regions of high gradient values. The map is expressed as a local weighted contrast:

$$W^k_{con}(i,j) = \sqrt{\sum_{p=-P}^{P} \sum_{q=-Q}^{Q} \omega_{p,q} \left[I^{gray}_k(i+p, j+q) - \mu_k(i,j) \right]^2},$$
$$\mu(i,j) = \sum_{p=-P}^{P} \sum_{q=-Q}^{Q} \omega_{p,q} I^{gray}_k(i+p, j+q), \tag{6}$$

where $i \in \{1, 2, ..., M\}, j \in \{1, 2, ..., N\}$ are spatial indices, M and N are image dimensions. $\omega = \{\omega_{p,q} | p = -P, ..., P, q = -Q, ..., Q\}$ is a 2D circularly symmetric Gaussian weighting function sampled out to 3 standard deviations ($P = Q = 3$) and rescaled to unit volume [17], and I^{gray}_k is the grayscale version of I_k.

Normalized weight maps are obtained to ensure that they sum to unity as follows:

$$\overline{W}^k = W^k / \sum_k W^k, \tag{7}$$

where $W^k = W^k_{sat} W^k_{chr} W^k_{sal} W^k_{fog} W^k_{lum} W^k_{con}$, and k is the index of I_k.

C. Multi-scale refinement

Multi-scale refinement [18] is used to achieve a halo-free defogged image. Each preprocessed image and corresponding normalized weight map are decomposed using a Laplacian pyramid, then they are blended to yield a fused pyramid

$$F_l = \sum_k G_l\{\overline{W}^k\} L_l\{I_k\}, \tag{8}$$

where l is the number of pyramid levels. In our experiment, l = 9 to eliminate fusion degradation. $G_l\{\cdot\}$ and $L_l\{\cdot\}$ mean the Gaussian and the Laplacian decomposition at pyramid levels

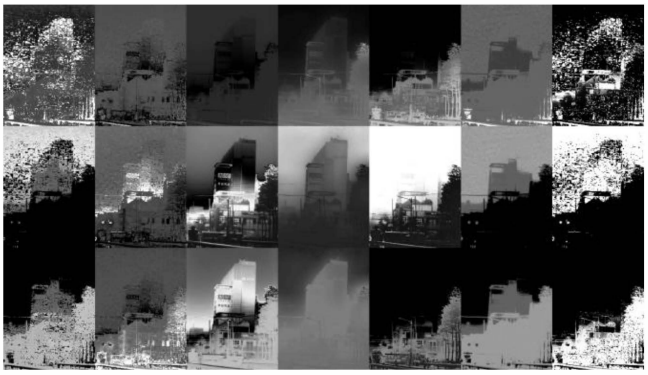


Figure 3. Weight maps. The first, second, and third rows are weight maps of the preprocessed images, I_1, I_2, and I_3, shown in Figure 2, respectively. Saturation, chrominance, saliency, perceptual fog density, luminance, contrast, and normalized weight maps are shown from left to right column.

l, respectively. Operations are executed successively for each pyramid layer, in a bottom-up manner. A defogged image J is achieved by the Laplacian pyramid reconstruction as follows,

$$J = \sum_l F_l \uparrow^n, \tag{9}$$

where $\uparrow^n$ is the upsampling operator with factor $n = 2^{l-1}$ [9].

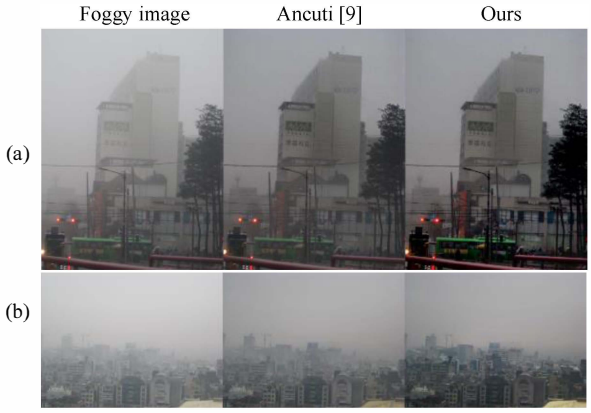
IV. Results

A large number of foggy images were tested to evaluate the performance of the proposed model. First, to explore the importance of fog aware statistical features that capture perceptual fog density in a fusion based defogging algorithm, we compared the results obtained using the method of Ancuti *et al*. [9] and ours on darker, denser foggy images in Figure 4. Results show that our model achieves better restoration of the contrast of edges and colors. Quantitative evaluation of defogged outputs was performed using the blind measure of Hautière *et al.* [19] and the perceptual fog density, D, of Choi *et al.* [12]. The metrics e, Σ, and r denote newly visible edges, the percentage of pixels that become black or white after defogging, and the mean ratio of the gradients at visible edges, respectively. Table I shows that the proposed model yields more naturalistic, clear edges after defogging while maintaining lower percentage of black or white pixels. D denotes that foggy images are more defogged in our model.

Figure 5 and Table II show results on standard defog test images for the models of Tan [4], Fattal [5], He *et al.* [6], Tarel *et al.* [8], Ancuti *et al.* [9], and ours. While He *et al.* [6], Ancuti *et al.* [9], and ours recover visible edges yielding positive values of the metric e, our model reduces perceptual fog density most significantly among these. Although Tan [4] achieves the most reduction of perceptual fog density, since it increases the local contrast too strongly, this method removes visible edges and has higher values of metric Σ and r. Results demonstrate that our model obtains comparable or better visibility enhancement than the compared models.

V. Conclusion and Future Work

We presented a referenceless perceptual image defogging model based on fog aware statistical features. The fog aware weight maps effectively filter the most visible areas of three preprocessed images and smoothly blend them via a multi-



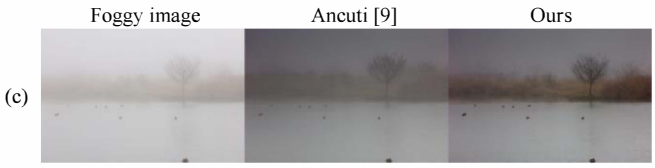


Figure 4. Comparison of defogged images on Ancuti *et al.* [9] and ours.

TABLE I. QUANTITATIVE COMPARISON OF DEFOGGED IMAGES SHOWN IN FIGURE 4 USING *e*, Σ, *r* OF HAUTIÈRE *et al.* [19] AND *D* OF CHOI *et al.* [12].

Foggy image		Ancuti *et al.* [9]				The proposed model			
	D	*e*	Σ	*r*	*D*	*e*	Σ	*r*	*D*
(a)	6.66	0.28	0.00	1.12	4.39	0.63	0.04	1.82	2.05
(b)	9.58	0.27	0.00	1.03	7.99	0.99	0.00	1.44	5.88
(c)	11.32	13.06	0.00	1.50	6.61	58.59	0.00	2.41	4.37

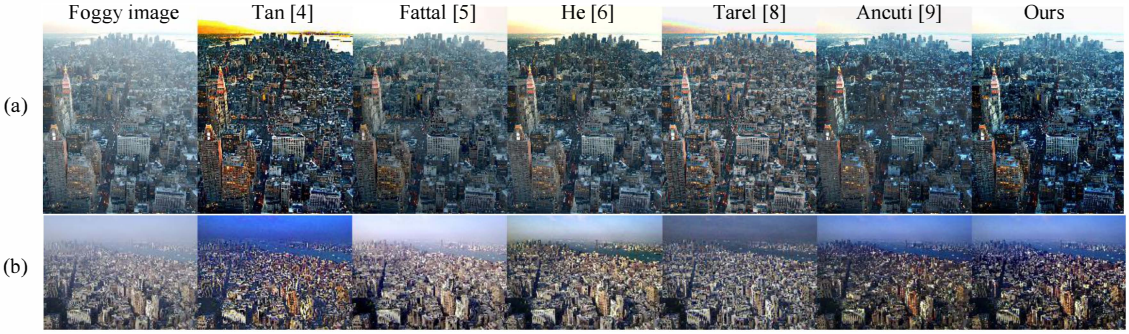


Figure 5. Comparison of defogged images on Tan [4], Fattal [5], He *et al.* [6], Tarel *et al.* [8], Ancuti *et al.* [9], and the proposed model.

TABLE II. QUANTITATIVE COMPARISON OF DEFOGGED IMAGES SHOWN IN FIGURE 5 USING *e*, Σ, *r* OF HAUTIÈRE *et al.* [19] AND *D* OF CHOI *et al.* [12].

Foggy image		Tan [4]				Fattal [5]				He [6]				Tarel [8]				Ancuti [9]				The proposed model			
	D	e	Σ	*r*	*D*	*e*	Σ	*r*	*D*	*e*	Σ	*r*	*D*	*e*	Σ	*r*	*D*	*e*	Σ	*r*	*D*	*e*	Σ	*r*	*D*
(a)	1.46	-0.09	0.72	2.57	0.36	-0.06	0.09	1.32	0.89	0.06	0.00	1.42	0.85	0.07	0.00	1.88	0.63	0.02	0.00	1.49	0.80	0.05	0.01	1.47	0.42
(b)	1.35	-0.10	1.28	2.29	0.34	-0.12	0.02	1.56	0.73	0.01	0.00	1.65	0.56	-0.01	0.00	1.87	0.52	0.12	0.00	1.54	0.57	0.01	0.00	1.44	0.39

scale Laplacian refinement. Results show that the proposed model achieves better performance for darker, denser foggy images as well as on standard defog test images. In future work, we plan to build a larger foggy image database and perform a human subjective test to understand the human perception of foggy images.

REFERENCES

[1] S. Narasimhan and S. Nayar, "Contrast restoration of weather degraded images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 6, pp. 713-724, Jun. 2003.

[2] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Instant dehazing of images using polarization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Dec. 2001, vol.1, pp. I-325- I-332.

[3] N. Hautière, J. -P. Tarel, J. Lavenant, and D. Aubert, "Automatic fog detection and estimation of visibility distance through use of an onboard camera," *Machine Vision and Applications*, vol. 17, no. 1, pp. 8-20, Apr. 2006.

[4] R. T. Tan, "Visibility in bad weather from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1-8.

[5] R. Fattal, "Single image dehazing," *ACM Trans. Graph., SIGGRAPH*, vol. 27, no. 3, p. 72, 2008.

[6] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1956-1963.

[7] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397-1409, Jun. 2013.

[8] J. -P. Tarel and N. Hautière, "Fast visibility restoration from a single color or gray level image," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep. - Oct. 2009, pp. 2201-2208.

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

[10] H. B. Mitchell, *Image Fusion: Theories, Techniques and Applications*. New York, NY, USA: Springer-Verlag, 2010.

[11] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 161-171, 2009.

[12] L. K. Choi, J. You, and A. C. Bovik, "Referenceless perceptual fog density prediction model," in *Proc. SPIE Human Vis. Electron. Imag.*, Feb. 2014, 9014-16.

[13] H. Koschmieder, "Theorie der horizontalen sichtweite," in *Beitrage zur Physik der Freien Atmosphare*. Munich, Germany: Keim & Nemnich, 1924.

[14] G. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Proc. 12th Color Imag. Conf.*, 2004, pp. 37-41.

[15] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1597-1604.

[16] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695-4708, Dec. 2013.

[17] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209-212, Mar. 2013.

[18] P. Burt and T. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532-540, Apr. 1983.

[19] N. Hautière, J. -P. Tarel, D. Aubert, and E. Dumont, "Blind contrast enhancement assessment by gradient ratioing at visible edges," *J. Image Anal. Stereol.*, vol. 27, no. 2, pp. 87-95, 2008.