ENHANCING THE PERCEPTION OF A HAZY VISUAL WORLD USING A SEE-THROUGH HEAD-MOUNTED DEVICE

Kai-En Lin, Kuang-Tsu Shih, and Homer Chen

National Taiwan University, Taipei, Taiwan

ABSTRACT

Human perception of the visual world in the presence of fog or haze usually suffers from loss of color vibrance and details. In this paper we propose a novel method to provide a see-through head-mounted display with image dehazing capability. The proposed method projects an auxiliary image from the head-mounted device to the user's retina. The introduction of the auxiliary image enhances the contrast and brightness of the perceived visual world. This method can be applied to tourism, military, and navigation to overcome poor visual condition.

Index Terms—See-through head-mounted display, augmented reality, dehazing, defogging, just-noticeable difference.

1. INTRODUCTION

See-through head-mounted devices (HMDs) have become more and more popular in recent years for applications such as education, entertainment, commerce, etc. This paper describes another interesting application, in which an HMD is used to enhance the user's vision by exploiting the see-through and augmented reality functions of the device. Specifically, the user's perception of a visual world that is fuzzy due to, for example, the presence of fog or smoke is enhanced by the projection of an auxiliary image onto the user's retina.

The user receives both the original scene light, which comes through the see-through HMD, and the device light from the projector, which is modulated by the auxiliary image of the scene. The addition of the modulated projection light results in an improved perception of the visual world that would otherwise appear hazy.

The color enhancement issue involved in this application is related to image dehazing, which has been a popular topic in the fields of computer vision and image processing for the last fifteen years. The goal there is to enhance an image of the visual world after it is captured by a camera, so the dehazing is performed as a post-processing step. The goal here is similar except that the dehazing is performed on-the-fly using a see-though HMD with a pico-projector and a camera; therefore, the color enhancement algorithm is different.

The main contribution of this work is that it proposes a novel method that allows a see-through head-mounted device to be used as a defog or dehaze device. It enhances the visual perception that would have been obscured otherwise for a user. As people heavily rely on the visual sense for object recognition, detection, and navigation, we believe the proposed method is useful for overcoming poor outdoor visual condition.

2. PREVIOUS WORK

The proposed method exploits previous work to create the auxiliary image. In this section, we give a brief review of the physical model used for image dehazing.

When a light ray travels through a murky atmosphere from a far distance, it attenuates along the way because of atmospheric scattering and absorption. Aerosols like haze, smoke, dust, and air pollutants all deflect light passing through them and attenuate the outgoing light rays. As a result, an image captured under such a weather condition would suffer from poor visual quality.

The main idea of image dehazing, or defogging, is to estimate the airlight [1] at different depths and remove the haze accordingly. Based on the groundwork laid down by Middleton [2] and McCartney [3], Nayar and Narasimhan [4] assumed homogeneous airlight and developed the following model for dehazing:

$$I(\mathrm{x}) = J(\mathrm{x})t(\mathrm{x}) + A(1 - t(\mathrm{x})), \tag{1}$$

where I denotes the observed scene luminance, x denotes a pixel location, J denotes the haze-free scene luminance, A denotes the airlight, and t denotes the medium transmittance that is related to depth d by

$$t(\mathrm{x}) = e^{-\beta d(\mathrm{x})}, \tag{2}$$

where β is the light scattering coefficient.

The problem described by (1) is ill-posed. We cannot solve for the haze-free image $J(\mathrm{x})$ unless the medium transmittance and airlight are given. However, we can make the problem solvable by using some appropriate assumptions or prior knowledge of the physical world. Assumptions such as maximizing local contrast [5], locally uncorrelated transmittance and surface shading [6], and dark channel [7] have been successfully applied to solve the problem. These

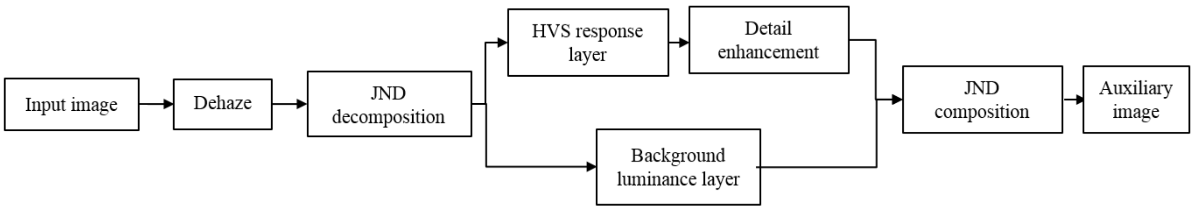

Fig. 1. Flow chart of the proposed method.

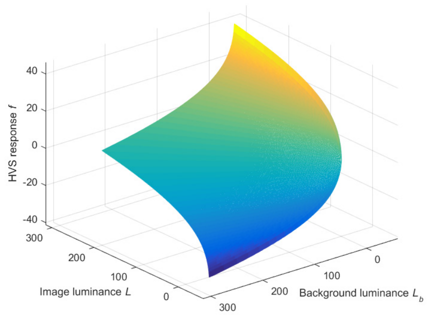

Fig. 2. The HVS response as a function of the input luminance and the background luminance.

methods can be applied to generate the auxiliary image in our work.

3. DEHAZING FOR AUGMENTED REALITY

3.1. Main Idea

Conventional image dehazing can freely adjust every pixel in the hazy image. However, for the application considered here, the scene light cannot be changed in the same way because it enters our eyes directly from the hazy scene. We propose to use an HMD to solve the problem because the device light projected from the HMD is controllable at the pixel level and offers a mechanism to manipulate light rays entering the user's retina. In this approach, the luminance P perceived by a user is a mixture of the scene light P_s and the device light P_d emitted from the projector,

$$P = P_s + P_d. \tag{3}$$

The main task is to generate an auxiliary image for the projector to improve the perceived brightness and contrast of the scene.

3.2. Assumptions

Two assumptions are made in this work. The first assumption is that the scene light and the device light are approximately at the same intensity level. This is a reasonable assumption because we can always adjust the device light intensity to match the scene light or use the dimming visor on the HMD to tone down the scene light if it is too bright. The second assumption is that the image of the scene capture by the camera of the HMD is the same as the one perceived by our eyes. Normally, the camera is installed on the perimeter of the glasses on the HMD, so there is a slight difference in viewing angle between the eye and the camera. This assumption implies that a high-quality calibration and interpolation algorithm exists.

3.3. JND Decomposition

We adopt the just-noticeable difference (JND) decomposition technique [8], [9] to extract details from the scene, because restoring details obscured by haze can enhance our perception of the hazy world. In particular, to match human visual perception, we decompose an image into two layers: the background luminance layer and the HVS response layer. The former characterizes local average luminance, while the latter characterizes the perceptual difference between a pixel and its neighbors.

The JND decomposition consists of two steps. In the first step, we apply a low-pass filter to the input image to compute the background luminance layer. In the second step, we compute the HVS response from the input image and the resulting background luminance. The HVS response can be represented by a non-linear bivariate function $f(\cdot,\cdot)$ shown in Fig. 2. The function is based on the JND concept in vision science [10], but it is independent of the input image. Note that for a fixed background luminance, the HVS response is a monotonic function of the image luminance. Therefore, if the background luminance and the HVS response are given, but the image luminance is to be determined, then we have

$$L = f^{-1}(H, L_b). \tag{4}$$

where L denotes the image luminance, L_b the background luminance, H the HVS response, and f^{-1} the inverse HVS response function. For more details on JND decomposition, please refer to the paper by Mantiuk et al. [9].

3.4. Generating the Auxiliary Image

To compute the auxiliary image, we capture an image of the hazy scene using a camera attached to the head-mounted display. Following the notation in (1), denote the captured hazy image by $I(\mathrm{x})$ and its haze-free counterpart by $J(\mathrm{x})$.

Fig. 3. The experimental setup. The camera is placed behind a pair of BT-200 AR glasses to capture images.

Our dehaze algorithm is based on the dark channel prior [7]. The dark channel of a pixel x in the image $I(\mathrm{x})$ is expressed as

$$\min_{\mathrm{y}\in\Omega(\mathrm{x})}\big(\min_{c\in\{\mathrm{r,g,b}\}} I^c(\mathrm{y})\big), \tag{5}$$

where $\Omega(\mathrm{x})$ denotes a local window centered at x and c a color channel of the image. The prior states that for a natural haze-free image, the dark channel is usually zero

$$\min_{\mathrm{y}\in\Omega(\mathrm{x})}\big(\min_{c\in\{\mathrm{r,g,b}\}} J^c(\mathrm{y})\big)\to 0. \tag{6}$$

Taking the dark channel of (1) yields

$$\min_{\mathrm{y}\in\Omega(\mathrm{x})}\big(\min_c I^c(\mathrm{x})\big)=\tilde{t}(\mathrm{x})\min_{\mathrm{y}\in\Omega(\mathrm{x})}\big(\min_c J^c(\mathrm{y})\big)+A^c(1-\tilde{t}(\mathrm{x})), \tag{7}$$

where $\tilde{t}(\mathrm{x})$ is the estimated medium transmittance (or transmission). Dividing both sides of (7) by A^c yields

$$\min_{\mathrm{y}\in\Omega(\mathrm{x})}\Big(\min_c \frac{I^c(\mathrm{x})}{A^c}\Big)=\tilde{t}(\mathrm{x})\min_{\mathrm{y}\in\Omega(\mathrm{x})}\Big(\min_c \frac{J^c(\mathrm{y})}{A^c}\Big)+(1-\tilde{t}(\mathrm{x})). \tag{8}$$

According to (6), the first term on the right-hand side of (8) equals zero. Therefore, the medium transmittance $\tilde{t}(\mathrm{x})$ is estimated as follows

$$\tilde{t}(\mathrm{x})=1-\omega\min_{\mathrm{y}\in\Omega(\mathrm{x})}\Big(\min_c \frac{I^c(\mathrm{x})}{A^c}\Big), \tag{9}$$

where the parameter ω is added to control the amount of haze removal. The estimated transmittance map is then refined by a guided filter [11] with the original input image as the guidance. The airlight A is estimated by averaging the values of the top 0.05% brightest pixels. A haze-free scene is then computed as follows:

$$J(\mathrm{x})=\frac{I(\mathrm{x})-A}{\max(t(\mathrm{x}),t_0)}+A. \tag{10}$$

Using the JND decomposition described in Sec. 3.3, the proposed method separates J(x) into the HVS response layer and the background luminance layer. The former contains visual details of the haze-free scene. Therefore, we can improve the visibility of these details effectively by adding a gain that is larger than 1 to the HVS response layer of J(x). Specifically, the output, which is the auxiliary image, is obtained by

$$I_a(\mathrm{x})=f^{-1}(\lambda f(J(\mathrm{x}),J_b(\mathrm{x})),J_b(\mathrm{x})), \tag{11}$$

where $\lambda>1$ is an adjustable parameter of the algorithm, f and f^{-1} are the non-linear functions described in Sec. 3.3, and J_b(x) is the background luminance layer of J(x).

4. EXPERIMENTAL RESULTS

Our choice of the see-through AR glasses is an Epson BT-200. As stated in Sec 3.2, we assume that a high-quality view interpolation algorithm exists. Therefore, we place the camera at the eye position in the experiment. The camera model we use is Canon EOS 7D Mark II. The zoom lens is adjusted to acquire image projected by the HMD as large as possible in order not to lose too much image resolution after image cropping. The dehaze algorithm is implemented in MATLAB using the dark channel prior [7]. Parameters for the dehaze algorithm are selected empirically.

In our experiments, we display a hazy image on an liquid-crystal display placed in front of the BT-200. Then, we capture the scene image and input it to the proposed method. The output, which is the auxiliary image, is projected onto the AR glasses. The proposed method is compared with a naïve method that only performs haze removal and uses the resulting image as the auxiliary image.

Test images from previous work and the D-HAZY dataset [12] are used here. All images shown are taken with the same exposure setting. Thus, the hazy scene images are dimmer than the combined view.

From Fig. 4 (b), we can see that the naïve method yields a fuzzy and blurry view, and that in Fig. 4 (c) the proposed method enhances details such as the balconies of the red buildings at the center and the orange buildings at a far distance of the scene. The periodic table in Fig. 4 (e) is hazy and the numbers are difficult to read. In contrast, numbers like “67” on the top of “Ho” in Fig. 4 (f) are recognizable, and the edges of the wood panel are relatively sharper. In Fig. 4 (i), it can also be seen that our method improves the details of objects such as the texture of the hat, the text on the cardboard box, and the sharpness of the plant.

5. CONCLUSION

In this paper, we have described a novel method that uses a see-through head-mounted device to enhance the visual perception of a hazy world. The method exploits the projector embedded in the device to enhance the view of a hazy scene at the pixel level by projecting an auxiliary image of the scene on to the retina of a user. The computation of the auxiliary image is based on JND decomposition, which allows us to extract haze-free details of the scene from the image data. We have demonstrated the advantage of integrating a see-through HMD with a dehazing software. Experimental results show that the proposed method yields a sharper and clearer view of a hazy scene.

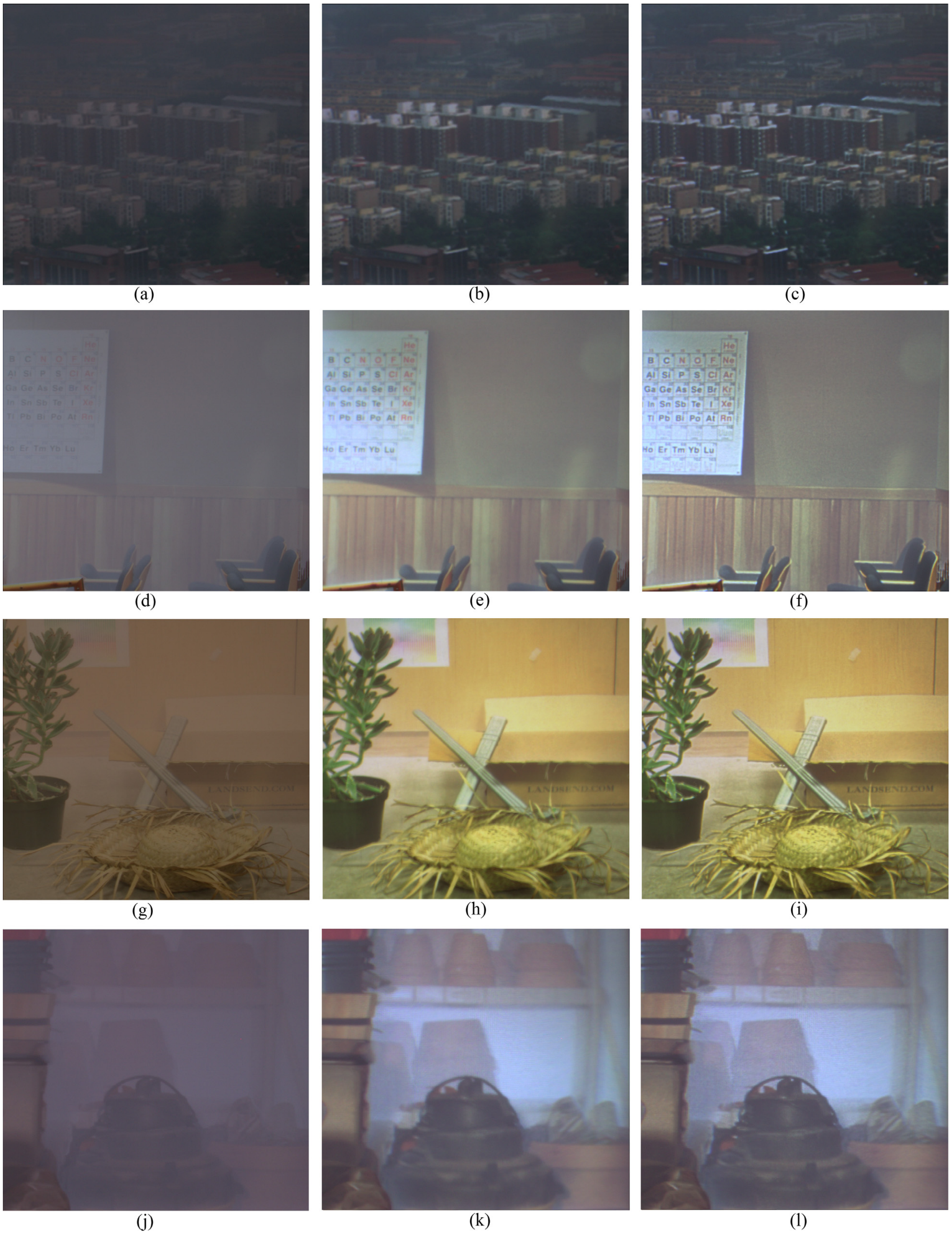

Fig. 4. From left to right: original hazy image, result of the naïve method, and result of the proposed method.

6. REFERENCES

[1] H. Koschmieder, “Theorie der horizontalen sichtweite,” in *Beitrage zur Physik der freien Atmosphere*., 1924, pp. 171-181.

[2] W. Middleton, *Vision through the atmosphere*. Toronto: University of Toronto Press, 1952.

[3] E. McCartney, *Optics of the atmosphere: scattering by molecules and particles*. New York, 1975.

[4] S.K. Nayar, and S.G. Narasimhan, “Vision in bad weather,” *Proc. of the Seventh IEEE Int'l Conf. Computer Vision*, vol. 2, pp. 820-827, 1999.

[5] Robby T. Tan, “Visibility in bad weather from a single image,” *IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.

[6] R. Fattal, “Single image dehazing,” *ACM SIGGRAPH*, vol. 27, no. 72, 2008.

[7] K. He, J. Sun, and X. Tang, “Single image haze removal using dark channel prior,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 2341-2353, 2010.

[8] T.-H. Huang, K.-T. Shih, S.-L. Yeh, and H. H. Chen, “Enhancement of backlight-scaled images,” *IEEE Trans. on Image Processing*, vol. 22, pp. 4587-4597, 2013.

[9] R. Mantiuk, K. Myszkowski, and H. P. Seidel, “A perceptual framework for contrast processing of high dynamic range images,” *ACM Trans. Appl. Perception*, vol. 3, no. 3, pp. 286-308, Jul. 2006.

[10] H. Wilson, “A transducer function for threshold and suprathreshold human vision,” *Biol. Cybern.*, vol. 38, no. 3, pp. 171-178, Oct. 1980.

[11] K. He, J. Sun, and X. Tang, “Guided image filtering,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 35, pp. 1397-1409, 2012.

[12] C. Ancuti, C. O. Ancuti, and C. D. Vleeschouwer, “D-HAZY: A dataset to evaluate quantitatively dehazing algorithms,” *IEEE Int'l Conf. Image Processing*, 2016.