International Journal of Engineering & Technology, 7 (2.12) (2018) 76-79

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Research of improved object detection with image dehazing

Hyochang Ahn [1] *, June-Hwan Lee [1], Han-Jin Cho [1]

[1] Dept. of Smart & PhotoVoltaic Convergence, Far East Univ., 27601, Republic of Korea,
**Corresponding author E-mail: 1youcu92@gmail.com*

Abstract

Background/Objectives: In this paper, we propose an improved object recognition method that can recover visibility by detecting degraded image due to factors such as weather, and to detect objects in the reconstructed image.
Methods/Statistical analysis: In this paper, we propose improved object detection and recognition method by visibility restoration by improving degraded image. First, fog removal based on a single image is performed and restored using a median filter instead of using the minimum value in the window. We also used adaptive feature point extraction method to extract feature points by removing unnecessary information such as noise from the improved image.
Findings: The performance of the proposed algorithm was evaluated by measuring the number of feature points before and after the proposed algorithm. Experimental results show that more feature points are extracted and matched than the case of recognition of the region of interest after reconstructing the image by removing the fog using the intermediate value filter in the image of degraded image by the proposed method.
Improvements/Applications: Applying the proposed method to a variety of vision - based intelligent systems such as smart cars will eliminate the bad conditions such as fog, which will give better performance.

Keywords: *Object Detection; Feature Descriptor; Feature Matching; Dark Channel; Image Dehazing*

1. Introduction

Detection of moving objects from video images has been the most basic and important field in the field of computer vision. However, most studies have been developed to provide optimal performance in a clean or laboratory-like environment that does not take into account environmental factors such as weather [1 2]. Most motion detection methods use a background difference technique using a fixed camera. The background difference is determined by determining the difference between the background created by the background model and the current input image [3 4]. However, the background is affected by time, weather, speed of movement, shadows caused by objects, etc.
Many images acquired in the outdoor environment are degraded in quality due to various external factors[5]. It is especially affected by weather such as fog, smoke, rain, snow. Such deterioration in image quality caused by the weather can cause a serious problem in the performance of the vision system or the monitoring system using the camera. Recently, a smart car system has installed a camera in a vehicle to detect collision or danger, and has evolved into an intelligent system that combines various vision technologies with a simple black-box system of a conventional recording system. However, there are environmental factors due to weather as the biggest obstacle to the development of this system [3-5]. In particular, fog causes scattering and attenuation of light to make the colors of surrounding areas all alike, making the color saturation very difficult, making it difficult to distinguish the shape of objects. Therefore, the performance of vision technologies to detect and track objects based on the color or motion of objects is poor [6-8]. Therefore, in this paper, we propose an improved object recognition method that can recover visibility by improving the degraded image due to weather factors and detect and track objects in the reconstructed image.

2. Related works

Recently, many methods for removing fog have been suggested by many researchers. Most fog removal methods have been studied to improve contrast using various statistical or physical properties [9-10]. Fog images have various atmospheric attenuation factors due to light absorption and scattering. The biggest feature of this image is the blurring of the scene due to decrease of visible distance and loss of high frequency component [5-9]. In addition, fog makes the color of the object similar to the fog in distance, and makes it difficult to distinguish the shape of the object by decreasing the saturation [4-5]. Therefore, early fog removal methods simply use the method of making the distinction between the background and the object clear by improving the contrast of the image. However, the simple contrast improvement method was not efficient. Therefore, many researchers have proposed various methods using multiple images or using external devices. These research methods can be roughly divided into two areas [3-4]. The first is based on contrast enhancement and the second is based on image quality restoration.
A representative method based on contrast enhancement is histogram smoothing [1-5]. In this method, color fidelity is not always maintained and a method of maximizing the contrast of the image is performed. The classic ways to improve contrast are un-sharp masking, gamma correction, and histogram smoothing. These methods reset the approximate contrast of the entire image. The main characteristic of the attenuation caused by the fog image is the local contrast reduction of the image with distance [2-3]. Therefore, these methods improve the image by maximizing the local contrast reduced by fog. A method based on image reconstruction improves

the contrast of the scene using a physical model to estimate the attenuation pattern of the image. It also uses additional information to get better results. These additional information is the distance of the scene, the correlation between shadow and transmission, empirical assumptions, and the nature of the appropriate scene.
The methods based on the initial image restoration use various additional information [4-5]. The effect of fog depends on the distance between the object and the camera [4-5]. That is, there is a difference in the attenuation amount according to the depth information. So we use multiple images under various fog conditions to calculate depth information. This method requires taking multiple pictures of the same scene under various weather and time conditions. If there is a difference in these conditions, the accuracy of the results will be very low [9-10]. However, it is not easy to acquire images of these conditions in real applications. Another method is to use polarized light. The method using polarized light is also one of the methods using multiple images. The difference is that more than two polarizing filters are used. The images acquired by the respective polarizing filters exhibit different degrees of polarization with respect to the fog. This method ensures that the direct delivery of the object is not polarized but only polarized to atmospheric light. However, this method is also very inconvenient to use in practical applications.
To overcome these drawbacks, many researchers start to consider fog removal methods by estimating depth information in a single image[9-10]. However, it is practically impossible to directly estimate depth information from a single image. Therefore, many researchers have proposed statistical characteristics of fog images, transmission methods using optical models, or methods of estimating atmospheric light.
In addition, there is a method using the feature points of the object in order to separate the background area from the reconstructed image and correctly extract and recognize the moving object[11,12]. In order to recognize objects, feature point extraction methods have been continuously studied so that objects can be represented well. (SIFT) and speeded up robust feature (SURF), which are typical algorithms for extracting feature points that characterize these objects[13,14]. The SIFT algorithm and the SURF algorithm have common purposes in terms of finding feature points, but they differ in performance[14-15]. The SURF algorithm uses the Hessian matrix of the detector to improve the speed of the SIFT algorithm. The SURF algorithm has faster computation speed, but has a disadvantage in that it has poor extraction performance compared to SIFT[15]. In this paper, we use the SURF algorithm for real - time use of SIFT algorithm and SURF algorithm. We propose a method that uses only depth information to increase the computation speed by reducing the number of feature points.

3. Proposed method

3.1. Image dehazing

It is difficult to distinguish the color and form of objects because it is difficult to ensure visibility when there is haze in the outdoors. In addition, it is difficult to apply various image recognition algorithms because haze causes the color and edge information of objects to decrease. Therefore, it is necessary to improve the visibility of the fog image and remove the fog in the image in order to apply various image processing algorithms.We use the fog modeling formula proposed by Narasimhan to remove the fog contained in a single image.

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

Where $I(x)$is the acquired image, A is the degree of brightness of the fog, and $t(x)$ is the amount of light transmitted to the camera without scattering. $J(x)$ is the brightness value of the original image. Therefore, fog removal is to obtain$A$, $t(x)$, and $J(x)$ from the input image $I(x)$. Since $t(x)$ has a value less than 1, the brightness of the acquired image can be seen as a combination of a clear image with no mist and a depth (amount) of haze.
When one pixel value of a dehazed image is checked, one of the RGB channels has a very dark value. That is, at least one value having a very small value exists in a certain area. The expression (2) is as follows.

$$J^{dark}(x) = min_{y\in\Omega(x)}\left(min_{c\in rgb}J^c(y)\right) \approx 0 \quad (2)$$

In equation (2), $J^{dark}(x)$ denotes the dark channel prior, J_c denotes the channel of the input image J, and Ω denotes a certain region centered on the point x. In the fog free image, we can see that the dark channel prior obtained from the above equation has a value of 0 in most areas except sky. In the case of a foggy image, the dark channel prior has a bright value due to the brightness of the fog. The dark channel prior has a brighter value as the fog becomes thicker as the proportion of Aincreases. Therefore, the density of fog can be predicted by using the dark channel prior obtained from equation (2), and the amount of transmission can be calculated using this.
The calculation of the amount of delivery is very simple. As mentioned earlier, in the dark channel prior to the fog free image, the RGB channel converges to 0, so $J^{dark}(x) \approx 0$ can be placed. Therefore, by transforming equation (3) as below, it is possible to remove $J(x)$ which is unknown information.

$$min_{y\in\Omega(x)}\left(min_{c\in rgb}\left(\frac{I^c(y)}{A^c}\right)\right) = t(x) \times$$
$$min_{y\in\Omega(x)}\left(min_{c\in rgb}\left(\frac{I^c(y)}{A^c}\right)\right) + 1 - t(x)$$

Since $J^{dark}(x) \approx 0$ can be set in equation (4), the part of $J(x)$ in the right part of equation (3) can be set to 0. Therefore, the transmission amount $t(x)$ is simply the value obtained by subtracting the normalized dark channel prior from 1 as the fog value.

$$t(x) = 1 - min_{y\in\Omega(x)}\left(min_{c\in rgb}\left(\frac{I^c(y)}{A^c}\right)\right) \quad (4)$$

However, since the computed value obtained from equation (4) is obtained by obtaining the minimum value in the local interval, the amount of transmission that does not coincide with the edge of the acquired image is obtained. Therefore, when this is used as it is, a phenomenon such as a halo effect occurs. To solve this problem, we applied the matting algorithm. Since the expression for matting-foreground and background is very similar to the modeling expression for foggy images, it is possible to match $t(x)$ to the edge information of the original image, assuming that $t(x)$ is a given trimap. Since the mattingtechnique uses a very large matrix, memory usage can be high and computation speed can be slow. Therefore, it is necessary to match the edge information of the image without using matting.Therefore, it is necessary to match the edge information of the image without using matting. In this paper, we use a medium filter instead of the minimum value in window ($N \times N$) to keep the large edge intact and smooth edges to smooth out the edge matching process and shorten the reconstruction time. However, there is a disadvantage in that the fine edges such as branches or leaves are softened and the halo effect is not completely removed. The equation to obtain the transfer amount map is expressed as Equation (5).

$$t(x) = 1 - \omega * med_{y\in\Omega(x)}\left(min_{c\in rgb}\left(\frac{I^c(y)}{A^c}\right)\right) \quad (5)$$

Result image acquisition can be obtained from equation (6). In the equation (6), $J(x)$gives the following equation.

$$J(x) = \frac{I(x)-A}{t(x)} + A \quad (6)$$

In the $t(x)$ and dark channel prior obtained from Eq. (7), the fog value Aselected from the brightest pixel values and $J(x)$from the

original image can be obtained. At this time, if the value of $t(x)$is too small, restored $J(x)$may have a very large noise, so that noise is suppressed by giving a threshold value as in the original paper. Therefore, the final restoration equation is as follows.

$$J(x) = \frac{I(x)-A}{max(t(x),t_0)} + A \qquad (t_0 = 0.1) \tag{7}$$

3.2. Object detection

SURF is an algorithm that quickly finds a point of interest that does not change even with changes such as size and rotation, and is suitable for image matching. In addition, SURF has an advantage of being able to perform partial image search without being affected by the preprocessing process. Since SURF is a feature calculated considering both large and small images, there is no change in the size of the image. The point of interest of SURF is obtained through the following steps. In order to increase the calculation speed of all rectangular areas, equation (8) calculates the integral image from the input image $I(x, y)$ through each point $X = (x, y)$.

$$I_{int}(x) = \sum_{i=0}^{i\le x}\sum_{j=0}^{i\le y} I(i,j) \tag{8}$$

For the scaled valueσ, equation (9) is applied to the Hessian matrix at a point $X = (x, y)$ in the image $I(x, y)$ to detect the candidates of interest.

$$H(X,\sigma) = \begin{bmatrix} L_{xx}(X,\sigma), L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma)\ L_{yy}(X,\sigma) \end{bmatrix} \tag{9}$$

$L_{xx}(X, \sigma)$, $L_{xy}(X, \sigma)$ and $L_{yy}(X, \sigma)$are the values obtained by the Gaussian second derivative.The approximated filter is represented by D_{xx}, D_{xy} and D_{yy} through the Gaussian second derivative. The Hessian determinant is calculated using equation (10).

$$det(H_{app}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \tag{10}$$

To compute feature points that are invariant to changes in size, the Hessian determinant is calculated by applying various filter sizes and scale ratios continuously to the image pyramid. The first size is a 9×9 filter, and the size value is assumed to be $\sigma = 1.2$. Next, 15×15, 21×21 and 27×27 filters can be applied. At this time, the size value of the 27×27 filter corresponds to

$$\sigma = 3 \times 1.2 = 3.6$$

After computing the Hessian determinant with the various filters, the point with locally localized local maxima becomes the point of interest k. That is, $3 \times 3 \times 3$ non-maximum suppressionsare performed on the image to locally limit the point of interest k. At this time, the maximum value of the Hessian determinant is interpolated according to each scaling factor.To be invariant to rotation, we need to know the direction at the point of interest. First, we use the Haar transform. After calculating the Harr transform for each pixel in the circle with a radius of 6σ at each point of interest, we weight the pixels in the 2σ distance from the point of interest to Gaussian. And the direction d in which the sum of all the weights is the largest is set as the direction k_d of the point of interest while the $3/\pi$ angle is rotated around the circle.

Even if there is a change in the rotation or size of spirituality, the descriptor does not change but uses the main direction and scale information of the minutiae to obtain a unique value, and rotates the image around the specific point and constructs a partial image with its size changed.Based on the main direction at the position of the center of the feature point, the image of a certain size region is rotated to be vertical based on the scale information, and the rotated image is converted into the size of 20×20 pixels to construct a partial image to extract the descriptor vector.

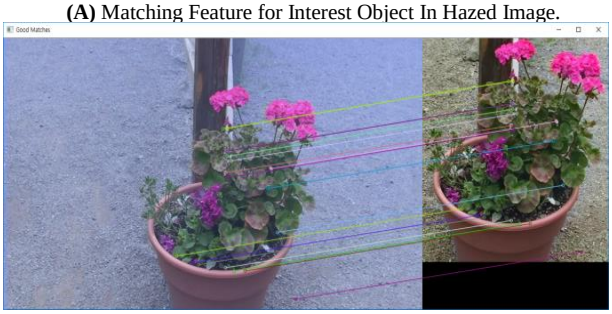
The partial image composed by image rotation and size conversion is divided into 16 areas, and the Haar wavelet response is calculated for each area to extract $\sum dx$, $\sum|dx|$, $\sum dy$ and$\sum|dy|$. Four vectors are extracted from 16 areas, and a total of 64 dimension descriptors are generated. Finally, a unit vector is normalized and stored as a descriptor vector for each region as a method for reducing the influence on the surrounding environment such as external light or illumination.

4. Results and discussion

In this paper, various image data are used for performance evaluation of improved object detection and recognition method through visibility restoration by improving image with degraded image quality. The system was also implemented using Visual Studio 2017 in a Windows 10 environment with specifications for i5 1.6GHz core and 8GB RAM. Experiments were performed by measuring the number of feature points before and after applying the proposed algorithm.

Fig. 1: to fig. 4: show the results of feature point extraction and matching for objects of interest in images with degraded image quality and visibility restored images using the proposed method.

(A) Matching Feature for Interest Object In Hazed Image.



(B) Matching Feature for Interest Object In Dehazed Image.

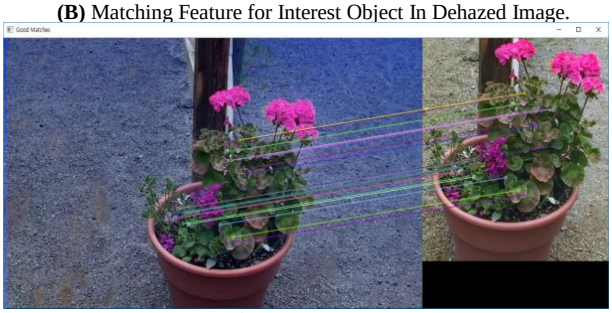


Fig. 1: Result of Pairwise Matching in Image 1.

(A) Matching Feature for Interest Object in Hazed Image.

(B) Matching Feature for Interest Object in Dehazed Image.

Fig. 2: Result of Pairwise Matching Features in Image 2.

(A) Matching Feature for Interest Object in Hazed Image

(B) Matching Feature for Interest Object in Dehazed Image.

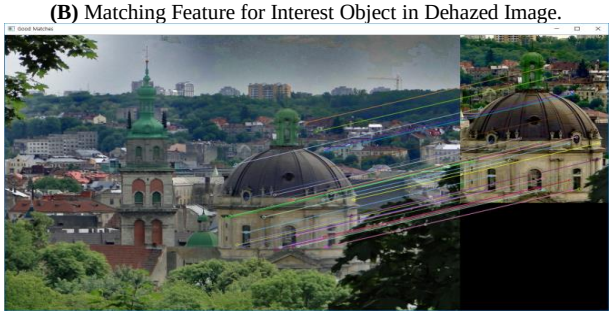


Fig. 3: Result of pairwise matching features in image 3.

(A) Matching Feature for Interest Object in Hazed Image.

(B) Matching Feature for Interest Object in Dehazed Image.

Fig. 4: Result of Pairwise Matching Features in Image 4.

The feature points corresponding to the feature points extracted from the object of interest are found, and the matching points are connected to each other to show the matching positions. Figure 5 shows the number of feature points before and after applying the proposed algorithm.

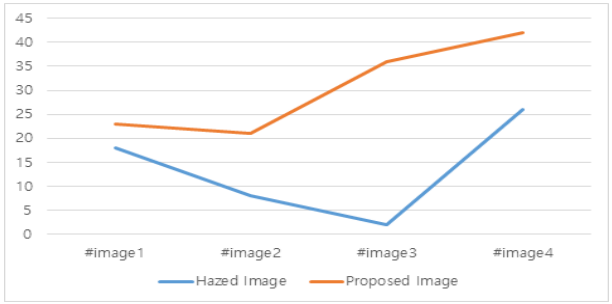


Fig. 5: Number of Pairwise Matching Features

As shown in the figure 5, when the image quality of the image is degraded due to factors such as fog, the number of feature points to be matched is reduced. When the number of matching minutiae is small, it is difficult to find an object because there is insufficient information to recognize the object. Therefore, after reconstructing the image by removing the fog using the intermediate value filter in the image with the degraded image quality by the proposed method, if the recognition of the region of interest is performed, accurate object recognition is possible because there are many matching feature points.

5. Conclusion

Experimental results show that the proposed method removes the fog by using the intermediate filter in the degraded image and restores the image. When the recognition of the region of interest is performed, more features are extracted and matched. Therefore, it was found that the proposed algorithm recognizes the objects quickly and efficiently. Future research needs to study the technique to detect collision and risk factors in various image based intelligent systems such as smart car.

References

[1] He, K., Sun, J., & Tang, X. Single image haze removal using dark channel prior. IEEE transactions on pattern analysis and machine intelligence, 2011, 33(12), pp.2341-2353.
[2] Gibson, K. B., Vo, D. T., & Nguyen, T. Q. An investigation of dehazing effects on image and video coding. IEEE transactions on image processing, 2012, 21(2), pp.662-673.
[3] Tan, R. T. Visibility in bad weather from a single image. In Computer Vision and Pattern Recognition, 2008. CVPR 2008. 2008, pp. 1-8.
[4] Gadnayak, K. K., Panda, P., & Panda, N. Haze Removal: An Approach Based on Saturation Component. In Intelligent Computing, Communication and Devices, 2015, pp.281-287.
[5] Lee, J. W., & Hong, S. H. Real-time Haze Removal Method using Brightness Transformation based on Atmospheric Scatter Coefficient Rate and Local Histogram Equalization. Journal of Korea Multimedia Society, 2016, 19(1), pp.10-21.
[6] Ahn, H., & Lee, Y. H. Performance analysis of object recognition and tracking for the use of surveillance system. Journal of Ambient Intelligence and Humanized Computing, . 2016, 7(5), pp.673-679.
[7] Viola, P., & Jones, M. Rapid object detection using a boosted cascade of simple features. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference, 2001, (Vol. 1, pp. I-I).
[8] Ahn, H. C., & Rhee, S. B. Fast Image Stitching Based on Improved SURF Algorithm using Meaningful Features. The KIPS Transactions: PartB, 2012, 19(2), pp.93-98.
[9] Gibson, K., Võ, D., & Nguyen, T. An investigation in dehazing compressed images and video. In OCEANS 2010 IEEE. 2010, September, pp. 1-8.
[10] Fattal, R. Single image dehazing. ACM transactions on graphics (TOG), 2008, 27(3), 72.
[11] Lowe, D. G. Distinctive image features from scale-invariant keypoints. International journal of computer vision, 2004. 60(2), pp.91-110.
[12] Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. Speeded-up robust features (SURF). Computer vision and image understanding, 2008, 110(3), pp.346-359.
[13] Brown, M., & Lowe, D. G. Invariant Features from Interest Point Groups. In BMVC. 2002, September, Vol. 4.
[14] Stommel, M., & Herzog, O. Binarising SIFT-descriptors to reduce the curse of dimensionality in histogram-based object recognition. Signal Processing, Image Processing and Pattern Recognition, 2009, pp.320-327.
[15] Mandle, P., & Pahadiya, B. An Advanced Technique of Image Matching Using SIFT and SURF. International Journal of Advanced Research in Computer and Communication Engineering, 2016, 5(5).nurses in korea. Diabetes Research and Clinical Practice 77, 199–204.
[16] McMahon GT, Gomes HE, Hohne SH, Hu TM, Levine BA & Conlin PR (2005), Web-based care management in patients with poorly controlled diabetes. *Diabetes Care* 28, 1624–1629.
[17] Thakurdesai PA, Kole PL & Pareek RP (2004), Evaluation of the quality and contents of diabetes mellitus patient education on Internet. *Patient Education and Counseling* 53, 309–313.