

Manhole Cover Detection from Natural Scene Based on Imaging Environment Perception

Haoting Liu^{1*}, Beibei Yan², Wei Wang², Xin Li³, and Zhenhui Guo³

¹ Beijing Engineering Research Center of Industrial Spectrum Imaging, School of Automation and Electrical Engineering, University of Science and Technology Beijing, No. 30 Xueyuan Road, Haidian District, Beijing 100083, China

[e-mail: liuhaoting@ustb.edu.cn]

² Department of R&D, Beijing Institute of Aerospace Control Device, No. 52 Yongding Road, Haidian District, Beijing 100854, China

[e-mail: bbyan_biacd@sina.com; wangwei_sdgd@126.com]

³ Jiuquan Satellite Launch Center, Jiuquan 732750, China

[e-mail: lixin_201984@protonmail.com; gzh9829@sohu.com]

*Corresponding author: Haoting Liu

*Received November 5, 2018; revised March 11, 2019; accepted March 17, 2019;
published October 31, 2019*

Abstract

A multi-rotor Unmanned Aerial Vehicle (UAV) system is developed to solve the manhole cover detection problem for the infrastructure maintenance in the suburbs of big city. The visible light sensor is employed to collect the ground image data and a series of image processing and machine learning methods are used to detect the manhole cover. First, the image enhancement technique is employed to improve the imaging effect of visible light camera. An imaging environment perception method is used to increase the computation robustness: the blind Image Quality Evaluation Metrics (IQEMs) are used to percept the imaging environment and select the images which have a high imaging definition for the following computation. Because of its excellent processing effect the adaptive Multiple Scale Retinex (MSR) is used to enhance the imaging quality. Second, the Single Shot multi-box Detector (SSD) method is utilized to identify the manhole cover for its stable processing effect. Third, the spatial coordinate of manhole cover is also estimated from the ground image. The practical applications have verified the outdoor environment adaptability of proposed algorithm and the target detection correctness of proposed system. The detection accuracy can reach 99% and the positioning accuracy is about 0.7 meters.

Keywords: Manhole cover detection, UAV, environment perception, image enhancement, deep learning

This research was supported by the National Natural Science Foundation of China under Grant No. 61501016 and the Fundamental Research Fund for the China Central Universities of USTB under grant No. FRF-BD-17-002A.

1. Introduction

With the fast development of artificial intelligence technology, the corresponding applications of smart city [1][2] meet great growth opportunities in recent years. The techniques of smart city use various methods to collect, transmit, and analyze data. The imaging sensor is one of basic facilities in big city [3]. Comparing with other sensors [4][5], the imaging sensor can acquire data with abundant contents, intuitionistic details, and vivid color; thus it can be utilized in many fields such as the intelligent surveillance or the public security protection, etc. For example, Chen X. et al. proposed a kind of cloud computation-based surveillance platform [6]. Many problems such as the weak tolerant ability or the low expansion performance of the traditional surveillance platform could be conquered in that platform. Hoang N.-D. et al. presented a system which could detect the surface crack in building structures automatically [7]. An improved Otsu method was utilized.

In this paper a multi-rotor Unmanned Aerial Vehicle (UAV) [8] is used to solve the manhole cover detection problem [9][10] in the suburbs of big city. Fig. 1 shows the manhole cover images captured by a multi-rotor UAV. Because of the stealing or the natural aging and damage, the city management department needs to monitor the working state of each manhole cover. However due to the staff shortage and the high cost of manpower, the management department has to seek a technical method to solve that problem. Obviously it is not necessary to build a video surveillance terminal for each manhole cover. Thus the multi-rotor UAV is a good choice for this problem [11]. After a primary test it can be found the application of multi-rotor UAV also has its own problems: first its imaging outputs are affected by the nature light seriously [12]; second the manhole cover detection ratio [13] is also limited after the primary test of the traditional circle detection-based methods [14].



Fig. 1. The manhole cover image samples captured by a multi-rotor UAV

In this paper, to realize a robust detection of manhole cover from the nature scene, a series of image processing and machine learning methods are utilized for the multi-rotor UAV-captured images. First, to improve the image output quality, a kind of environment perception-based image enhancement algorithm is proposed. The blind Image Quality Evaluation Metrics (IQEMs) are used to assess the imaging environment and select the image whose quality is qualified for the following processing. The blind IQEMs include [15] the image brightness degree, the image region contrast degree, and the image edge blur degree. Then a kind of adaptive Multiple Scale Retinex (MSR) method [16] is used to improve the output effect of the selected image. Second, the deep learning-based tool [17] is used to detect the manhole cover. The Single Shot multi-box Detector (SSD) method [18] is utilized here. Third, the ground coordinate [19] of manhole cover is also estimated from the ground image.

The main contributions of this paper include: first, an image enhancement algorithm which utilizes the blind IQEMs as the environment perception feedback is proposed. The application scopes of blind IQEMs are expanded; and it can improve the algorithm robustness for the outdoor application effectively. Second a practical UAV-based system is developed. The adaptive MSR and the SSD are also used to detect the manhole cover.

In the following sections, first the problem formulation and the system design method are presented. Second, the proposed processing algorithms are introduced. Third, some experiments and discussions are given.

2. Problem Formulation and System Design

The plane graph of the suburbs of a big city is shown in **Fig. 2**. It is generated by Google Earth. The crosses in orange color stamp the check points of manhole cover in that area. Fifty check points can be found in that map. The lines in yellow color mark the path that a man needs to walk in order to check the state of each manhole cover. The line in white color shows the crow-fly distance between the first check point and the fifteenth check point. The manhole covers locate in different terrains of this area including grass land, soil land and cement land. The total length of the yellow line is about 1.5 kilometers which will cost a man about 30 minutes to finish one way check travel. The length of the white line is about 1.0 kilometer. In the past, the management department asked the staff to take photo of each manhole cover during the check procedure; and one problem was different staffs used their personal photography habits to collect data [20]. Thus it is necessary to use a machine like the UAV to implement that task so that the manpower can be saved and the photography outputs can be standardized.



Fig. 2. The plane graph of the suburbs of a big city

A kind of multi-rotor UAV is developed in this paper. This system has four rotors. A visible light camera is fixed in its bottom. When it works, the real-time visible light image can

be transmitted back to the ground station [21]. Its average hover time is about 18 minutes; its proposed maximum flight height is about 50 meters; its maximum flight distance away from the ground station is about 2.0 kilometers; and its working temperature locates from -10°C to 40°C . When a user operates it, first its flight path [22] can be set by the Global Positioning System (GPS) points which are illustrated in the ground station. The ground station is installed in a PC. Second, the UAV can implement the flight mission by traversing each GPS point, capturing and transmitting the visible image back to the ground station. The ground station can save these data to PC. Third, the image processing task can be accomplished by the PC in real time; and the flight state parameters also can be shown to user.

3. Proposed Manhole Cover Detection Algorithm

3.1 The Algorithm Computational Flow Chart

Fig. 3 shows the proposed computational flow chart of manhole cover detection. The main problem in this algorithm is to develop a robust computation method to perform the manhole cover detection task in the nature scene. Thus the environment adaptability and the manhole cover detection ratio should be considered carefully. In **Fig. 3**, when the visible light image sequences are transmitted back, first the blind IQEMs are computed for each frame. Only the frame whose image quality is qualified can be accepted for the further processing. Second, a kind of adaptive MSR-based image enhancement method is employed. This algorithm can improve the imaging effect to some extent. Third, the enhanced image will be processed by the deep learning-based method. Finally, the coordinates of manhole cover will be estimated from the ground image. The geometrical optics-based ground resolution estimation method [23] is employed here.

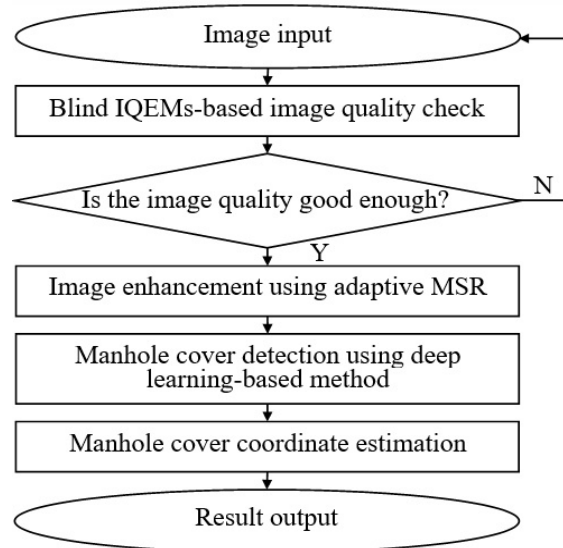


Fig. 3. The computational flow chart of proposed algorithm

3.2 The Environment Perception-based Image Enhancement

It is well known the image processing of UAV system is one of most difficult application cases for the algorithm design. Some algorithms will lose effectiveness or even fail due to the

UAV attitude change-caused environment light variety. To overcome that problem, the blind IQEMs are used to assess the nature environment here. A well designed blind IQEM can be independent to the image content. It can represent the inherent attributions of an image objectively. The image brightness degree, the image region contrast degree, and the image edge blur degree are computed here. Equations (1), (2), and (3) show their computation methods, respectively. Many blind IQEMs have been proposed in recent years [24], however after some practical tests it can be found that our proposed metrics are fit for engineering computation [25] better; i.e. their processing speeds are comparable fast and their processing effects are acceptable for this application.

$$M_{IBD} = \frac{1}{N_1} \left\{ \sum_{n_1=1}^{N_1} \left[\sum_{n_2=0}^{255} h_{n_1 n_2} \times (n_2)^s \right] \right\} \quad (1)$$

$$M_{IRCD} = \frac{1}{N_2} \sum_{k=1}^{N_2} \frac{I_k^{\max} - I_k^{\min}}{I_k^{\max} + I_k^{\min}} \quad (2)$$

$$M_{IEBD} = \max_{I \in \Theta} \left\{ \arctan \left[\frac{I(i_1, j_1) - I(i_2, j_2)}{W_{12}} \right] \right\} \quad (3)$$

where $h_{n_1 n_2}$ is the pixel quantity of the grey values n_2 in the histogram of the n_1 th image block; s is a parameter, $s=3$; N_1 is the number of sample blocks, $N_1=50$, the size of sample block is 100; I_k^{\max} and I_k^{\min} are the maximum and the minimum gray values of the k^{th} image block; N_2 is the number of sample block, $N_2=100$; Θ is the set of image block, $I(i_1, j_1)$ and $I(i_2, j_2)$ represent the grey values of the first and the second image blocks; W_{12} is the width of the edge-spread points (i_1, j_1) and (i_2, j_2) .

After the IQEMs are computed, only the image whose quality is qualified is saved for the further processing. Here the qualified image means its brightness is moderate, its contrast is distinct, and its edge is clear. A kind of adaptive MSR is used to enhance the image effect. Equations (4) to (8) show the computation method of it. Obviously, equations (4) and (5) are the calculation methods of the Single Scale Retinex (SSR); while equation (8) is the computation method of MSR. When performing the adaptive computation, first three SSRs with different scale factors ε_1 , ε_2 and ε_3 are computed. Second, the weight l_3 of SSR is set by user's experiences; in this paper $l_3=1/3$. Third, l_1 and l_2 are computed by equations (6) and (7); then the MSR is calculated by (8). From the computation steps above it can be seen this method can enhance image adaptively because the estimation of l_1 and l_2 have the adaptive mechanism. To improve the adaptive processing ability further, the neural network-based method [15] can be used in future.

$$SSR_k(i, j) = \log \left\{ \frac{I(i, j)}{[I(i, j) * G_k(i, j)]} \right\} \quad (4)$$

$$G_k(i, j) = \frac{1}{2\pi\varepsilon_k^2} \exp \left[-\frac{(i^2 + j^2)}{2\varepsilon_k^2} \right] \quad (5)$$

$$l_1 = \frac{\text{abs}\{\text{mean}[SSR_1(I)]\}}{\text{abs}\{\text{mean}[SSR_2(I)]\} + \text{abs}\{\text{mean}[SSR_1(I)]\}} \times (1 - l_3) \quad (6)$$

$$l_2 = \frac{\text{abs}\{\text{mean}[SSR_2(I)]\}}{\text{abs}\{\text{mean}[SSR_2(I)]\} + \text{abs}\{\text{mean}[SSR_1(I)]\}} \times (1 - l_3) \quad (7)$$

$$MSR(i, j) = \sum_{k=1}^3 l_k \times SSR_k(i, j) \quad (8)$$

where l_k is the weight, ε_k is the scale factor, $k=1, 2, 3$; $I(i, j)$ is the input image; $G_k(i, j)$ is a low pass Gaussian filter; the symbol “*” means the convolution computation; $SSR_1(I)$ and $SSR_2(I)$ mean to compute the SSRs of image I when the scale factors are ε_1 and ε_2 , respectively; $abs(*)$ and $mean(*)$ mean to calculate the modulus and the mean, respectively.

3.3 The Manhole Cover Detection Using Deep Learning Architecture

The deep learning architecture is a successful application of the complex neural network. Currently, the basic theory of deep learning is not perfect; however its successful applications [26] have covered its shortcomings, such as the time-consuming training, or the complex network parameters tuning, etc. The SSD is a kind of application of the deep learning network. This approach comes from the feed-forward Convolutional Neural Network (CNN). The SSD has three characters. First, it can implement the multi-scale feature map detection. Second, it utilizes the convolutional predictors for detection. Third, it employs the default boxes and the aspect ratios to discretize the space of the possible output box shapes efficiently. The computational method of its overall objective loss function is shown in (9). By using the computation mechanism above, its computation speed is very fast. After a primary experiment test for our application, it can be found the SSD can carry out the target detection task almost in real time [27].

$$L(x, c, l, g) = \frac{1}{N} [L_{conf}(x, c) + \alpha L_{loc}(x, l, g)] \quad (9)$$

where N is the number of matched default boxes; $L_{conf}(x, c)$ is the classification loss function; and $L_{loc}(x, l, g)$ is the location regression function. Symbol x is an indicator for matching the default box to the ground truth box; symbol c is a probability parameter; symbol l means the predicted box, symbol g indicates the ground truth box; and symbol α is a weight.

3.4 The Coordinate Estimation of Manhole Cover from Aerial Image

The coordinate estimation of manhole cover is necessary because the city management department needs to confirm the UAV-based check processes are correct. As we have stated above, the familiar terrains in suburbs of big city include mountain, river, or forest. These terrains will create complex wind near the ground. As a result the flight stability of the mini-type UAV will be influenced by the wind seriously. When performing the coordinate estimation of manhole cover, first the lens of camera is controlled to be parallel to the ground. Second, the electro-optical stabilized platform will tune the attitude of camera in real time so that its imaging outputs are smooth. Third, if the GPS information of UAV, the heading angle, and the offset angle between the electro-optical stabilized platform and the UAV are all known, the distance offset [28] between the manhole cover and the imaging center can be computed by (10) to (13).

$$R_x = [h \times \tan(\phi_x/2)]/W \quad (10)$$

$$R_y = [h \times \tan(\phi_y/2)]/H \quad (11)$$

$$Offset_x = R_x \times Offset_x \quad (12)$$

$$Offset_y = R_y \times Offset_y \quad (13)$$

where R_x and R_y are the ground resolutions in the vertical and the horizontal directions; h is the flight height; φ_x and φ_y are the view angles in the vertical and the horizontal directions; W and H are the camera resolutions in the vertical and the horizontal directions; $Offset_x$ and $Offset_y$ are the distance offsets between the manhole cover and the imaging center (it can be approximated by the geometrical center of multi-rotor UAV); $Offset_x$ and $Offset_y$ are the pixel offsets between the imaging manhole cover and the imaging center.

4. Experiments and Discussions

To test the correctness of proposed system and method, a series of simulation and actual UAV-based test experiments are carried out. The experiment data are all captured by a multi-rotor UAV from the nature scene. The simulation and test programs are implemented by C code.

4.1 The UAV System and Its Ground Software

A kind of four-rotor UAV system is developed by us. Fig. 4 shows the photos of the corresponding hardware and software systems; and Table 1 gives out the performance parameters of the visible light camera. In Fig. 4, (a) is the photo of an UAV system; (b) shows the photo of an optoelectronic pod (including the electro-optical stabilized platform and the visible light camera); (c) and (d) show the photos of the software interface. From Fig. 4 (c) it can be seen that the flight path can be set by the GPS coordinates in the ground software. The user can accomplish that task just by utilizing the mouse device. It can also illustrate the flight state of UAV including the flight speed, the flight direction, or the flight height, etc. Image (d) shows the software interface of manhole cover detection. If the influence of wind near the ground is small and the map precision is high enough, the UAV can position and hover just above a manhole cover; as a result the manhole cover will appear in the middle of image center. In other cases the manhole cover will not locate in that position.

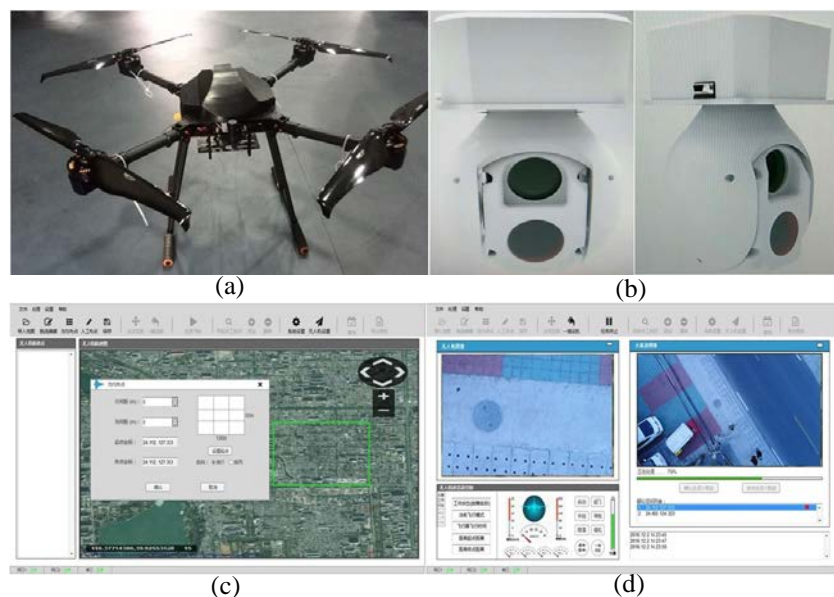


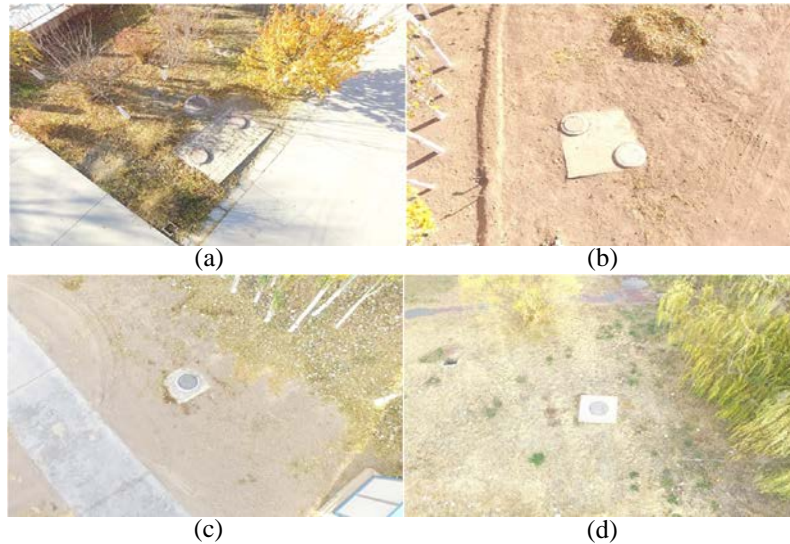
Fig. 4. The photos of multi-rotor UAV system and its ground soft-ware interfaces

Table 1. The performance parameters of the visible light camera

Sensor type	Sensor size (μm)	Image resolution	Focus range (mm)	Frame rate (Hz)	Maximum aperture (mm)
CCD	2.8	1920×1080	4.7-94.0	30	5.6
Minimum illuminance (lux)	Field angel ($^{\circ}$)	Automotive focus	Automotive gain	Wite balance	Data type
0.05	3.3-59.5	on	on	on	H.264

4.2 The Evaluations of Image Enhancement Algorithm

The image quality evaluation solves the environment perception computation in this paper. **Fig. 5** shows some ground image samples. In **Fig. 5**, (a) and (b) contain the multiple manhole covers in the same scenes; where (a) is affected by the shadows and (b) is influenced by the ground colors. Images (c) and (d) only have one manhole cover in their image contents. **Table 2** shows the image quality computation results of **Fig. 5** and the valid regions of the corresponding IQEMs. In **Table 2**, the valid distribution region means the corresponding image quality is qualified if the computed IQEMs locate in these regions. These regions can be defined by the system users. Finally, the quality check rule can be set by: regarding an image, if all its image quality evaluation results can locate in their respective valid regions then it will be accepted for the further processing. Obviously, from **Table 2**, it can be seen that **Fig. 5** (a) and (d) can pass the quality check.

**Fig. 5.** The image samples captured by a multi-rotor UAV**Table 2.** The image quality evaluation results of **Fig. 5** and the valid distribution regions of the corresponding IQEMs

Item	M_{IBD}	M_{IRCD}	M_{IEBD}
(a)	177.2748	3.6182	17.7356
(b)	191.3278	2.7098	16.5846
(c)	198.7259	2.7272	16.3151

(d)	191.6501	2.8795	18.0662
Valid region	[125.0, 195.0]	[2.8, 4.2]	[11.5, 20.5]
Description	The larger the brighter of image	The larger the better of image contrast	The larger the clearer of image edge

In our past research work [29], we have found the image whose imaging quality is high will behave a better processing effect when carrying out the following image feature computations. For example, the image with high quality will get more feature points or image edge details after computation; while the image with low imaging quality will contain a poor image features. Thus the image enhancement processing is necessary in this paper. Fig. 6 shows some processing results of the image enhancement experiment. In Fig. 6, (a) and (c) are the original images; while (b) and (d) are the processing results. Table 3 shows the image quality evaluation results of Fig. 6. From Tables 2 and 3, it can be seen images (a) and (c) cannot pass the image quality check at first; while after the enhancement processing their evaluation results turn to become valid. After some computations, it also can be found most of the image quality evaluation results of these UAV-captured images can be improved. These results can show the effectiveness of proposed enhancement method to some extent.

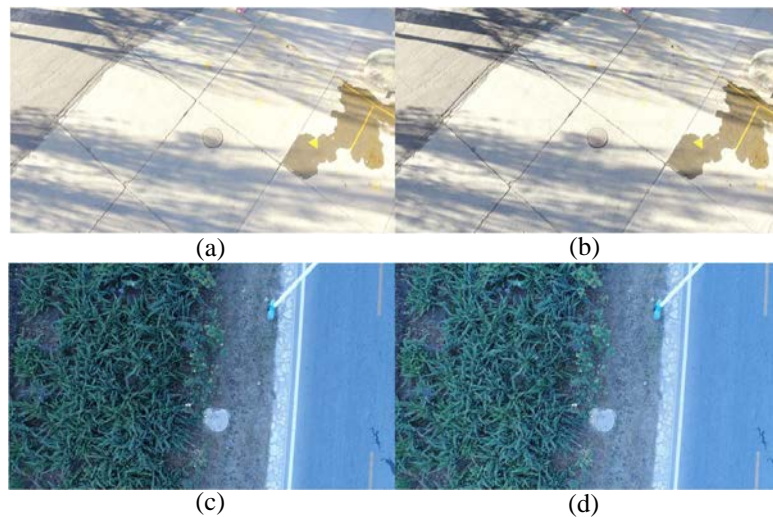


Fig. 6. The experiment results of image enhancement

Table 3. The image quality evaluation results of Fig. 6

Class	M_{IBD}	M_{IRCD}	M_{IEBD}
(a)	197.1287	3.1915	10.5330
(b)	187.7610	3.3387	12.4407
(c)	117.7142	3.7967	16.3924
(d)	130.6340	3.8057	16.7324

4.3 The Evaluations of Manhole Cover Detection Algorithm

The SSD is employed here due to its good computational performance for target detection. When using SSD, the initial position of manhole cover can be marked by hand firstly. A boundary rectangle is used to mark the position of manhole cover in our program. The original image data amount is 2544; and after translation, flipping, adding-noises, and size-cutting, the

data amount can reach more than 15000; and after hand marking this image dataset can be used to train the network. The amount of test dataset is 1000. The following open source code is used in our system because it uses the transfer learning mechanism for network training: <https://github.com/weiliu89/caffe/tree/ssd>. Finally, after training the SSD can be used for manhole cover detection. When evaluating the processing effect of SSD, the equations (14) and (15) are used [30]. Here the index Pr means the classification precision while the index Re is the recall rate. Table 4 shows the computation results of them. Fig. 7 shows some computation results of manhole cover detection. In Fig. 7, (a), (b), and (c) are captured in the morning of a sunny day; while (d), (e), and (f) are recorded at the noon of a sunny day. The red rectangles mark the detection results of the manhole covers. Obviously, (a) suffers from a low contrast problem; (b) is affected by the environment light; (c) is influenced by the grassland; (d) is disturbed by the shadow; (e) and (f) are mixed by the ground which has the similar color with the manhole cover. From Fig. 7 it can be seen that the SSD can perform the manhole cover recognition task well.

$$Pr = \frac{TP}{TP + FP} \quad (14)$$

$$Re = \frac{TP}{TP + FN} \quad (15)$$

where TP means the true positive, i.e., the number of targets correctly classified; TN means the true negative, i.e., the number of nontargets correctly classified; FP means the false positive, i.e., the number of nontargets classified as targets; FN means the false negative, i.e. the number of targets is classified as nontargets.

Table 4. The evaluation experiment results of SSD method

Item	Pr	Re
Result	0.997	0.994

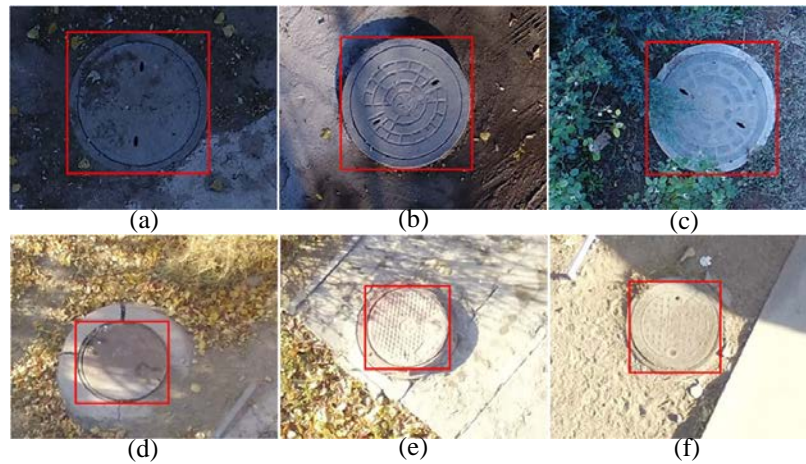


Fig. 7. The manhole cover recognition results using deep learning-based method

Many deep learning-based methods have been developed for target detection in recent years, such as YOLO [31], R-CNN [32], Resnet [33], Fast-CNN [34], and Faster-CNN [35] etc. Regarding the methods mentioned above, the R-CNN and the Fast-CNN are the early

versions of Faster-CNN; and their processing efficiencies are comparibale low. The Resnet is always used as a pre-training network for the target detection task. As a result we only make the evaluations among YOLO, Faster-CNN, and SSD in this paper. First, the performance analyses between YOLO and Faster-CNN are made. The pre-training model of Faster-CNN is VGG16, the development environment of it is windows 10, tensorflow, cuda toolkit 9.0, cudnn 7.1 and python 3.5; the pre-training model of YOLO is darknet53.conv.74, the development enviroment of it is windows 10, ubuntu 16.04, cuda toolkit 8.0, cudnn5.1 and python 3.5. The Graphic Processing Unit (GPU) of the experimental system is GTX 1080 8G. The dataset in VOC2007 (<http://host.robots.ox.ac.uk/pascal/VOC/voc2007/>) is used for the evaluation experiment. After computation it can be seen that even with the less iteration times, the mean average precision of YOLO is better than that of Faster-CNN which are 0.6423 and 0.6211, respectively. And the processing speed of YOLO is faster which almost can comform to the UAV real time application. The processing speed of YOLO is better than 10 frames per second. Second, a comparison between YOLO and SSD is also made. The dataset of manhole cover is utilized. After a computation it can be seen the SSD has a better processing effect; the mean average precisions of YOLO and SSD are 0.6541 and 0.712, respectively. The processing speed of SSD is faster than that of YOLO. As a result the SSD is selected in this paper.

To show the necessity of environment perception processing, an evaluation experiment is carried out. This experiment compares the parameters Pr and Re between the images which use the proposed enhancement method and the images which do not use it. **Fig. 8** shows the test image samples. In **Fig. 8**, (a) is the original manhole cover images; (b) is the enhanced ones. From **Fig. 8** it can be seen the image quality of (b) is better than that of (a). Then we can use the data like (a) and (b) to test SSD. **Table 5** shows the detection effect and the running speed comparisons between the enhanced images and the non-enhanced ones. From **Table 5** it can be seen the enhanced images will get a better detection effect. Regarding the processing speed, because the city management department thinks both the 14 frames per second and the 18 frames per second are acceptable for this application; as a result, when assessing the system performance, it can be thought the higher the detection ratio is, the better the system performance would be.

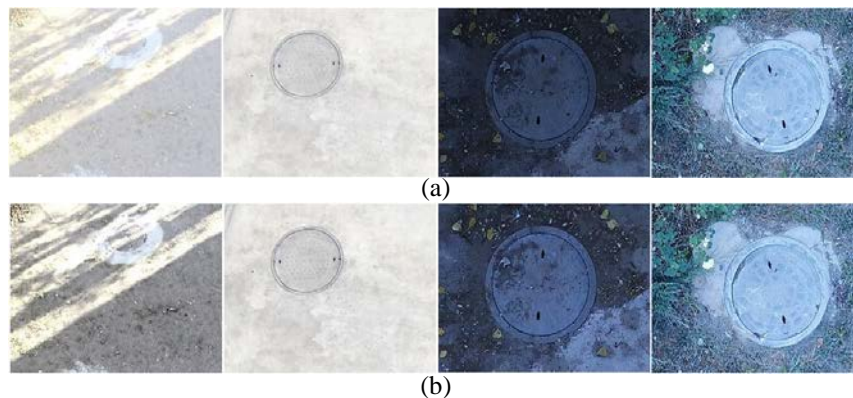


Fig. 8. The image data of the of environment perception necessity test experiment

Table 5. The detection effect and the running speed comparisons between the enhanced and the non-enhanced images

Item	Pr	Re	Running speed
Non-enhanced images (a)	0.991	0.990	≈18 frames/second
Enhanced images (b)	0.997	0.994	≈14 frames/second

To evaluate the computation effect of proposed method further, a kind of Hough circle detection-based method [36] is compared here. The processing steps of this experiment are: first our proposed image enhancement computation is performed. Second, the binary image segmentation is computed to get the initial image edge points. Third, the fast Hough circle detection is implemented. The random edge point selection method is employed here. Finally, the position of the manhole cover can be estimated by the geometrical constraint of a circle. **Table 6** shows the detection effect comparisons between the Hough circle detection-based method and our proposed method. From **Table 6** it can be seen our proposed method has a better processing effect. Because these methods above can all conform to the practical application requirement and implement the computation in real time (e.g. their processing speed can be large than 14 frame per second); it is not necessary to assess their processing speed in this paper any more.

Table 6. The detection effect comparisons between the fast Hough-based method and the SSD method

Item	<i>Pr</i>	<i>Re</i>
Hough circle detection-based method	0.984	0.976
Proposed method	0.997	0.994

4.4 The Evaluations of Manhole Cover Positioning Method

If the wind near the ground is strong, the UAV will deviate from its given GPS point; thus it is necessary to assess its positioning precision. When performing the positioning precision estimation experiment, first the user will use a portable differential GPS system to measure the coordinate of each manhole cover, e.g., (x_{p_i}, y_{p_i}) , where $i=1, 2, \dots, 15$. Second, the UAV is controlled to traverse all the GPS points. The GPS points are selected just in the spatial locations of manhole covers. Third, because of the near ground wind, the UAV will keep on floating in air even it arrives at that point and implements the hovering flight. Then the portable GPS system can be used again to measure the spatial position under UAV, e.g., (x_{uav_i}, y_{uav_i}) , where $i=1, 2, \dots, 15$. Finally the distance difference between UAV and manhole cover can be estimated by (16); the final error can be estimated by (17). After a series of tests the positioning precision of UAV is about 0.7 meters. Obviously, this error comes from the precision of Google Earth map, the influence of near ground wind and the control ability of UAV.

$$E_k = \frac{1}{15} \sum_{i=1}^{15} \sqrt{(x_{p_i} - x_{uav_i})^2 + (y_{p_i} - y_{uav_i})^2} \quad (16)$$

$$E = \frac{1}{K} \sum_{k=1}^K E_k \quad (17)$$

where E_k is the error estimation of one time flight; E is the final error estimation; K is the flight times, $K \geq 10$.

4.5 Discussions

With the fast development of information technology, the multi-rotor UAV has been utilized widely in recent years. The multi-rotor UAV can replace people to implement complex task that people cannot perform conveniently. The positive side of multi-rotor UAV application is

apparent; however, its application also has lots of shortcomings. The first one is its flight time is short. This is enslaved to the development state of battery technique of the mini-type UAV. This fact also indicates that the multi-rotor UAV must have a data processing algorithm with high efficiency; otherwise it will not have too much time to stay in air for information collection and processing. The second problem is the information processing ability of multi-rotor UAV is still limited. The multi-rotor UAV expands the activity fields of people; however the corresponding data processing methods, such as the target detection, the target recognition, and the target tracking, still need to be researched. To solve that problem well, one method is to develop some new high efficiency image processing algorithms; and the other measurement is to fall back on the intelligent hardware technique.

The nature environment creates great influences on the imaging quality, especially for the application case of multi-rotor UAV with visible light camera. The sun light, the haze, the fog, and even the surface features of target will affect the image recognition algorithm. To decrease that negative effect, the adaptive processing mechanism should be considered. In our past research work, the basic principle of the adaptive processing mechanism [15] can be summarized as: on one hand, sampling state data from the environment; on the other hand, implementing the data processing by the consideration of sampling results. However, the difficulty of that principle above is how to sample the environment characters from the environment. To be more specific, regarding the detection algorithm design of manhole cover, this problem becomes how to compute some kinds of features to describe the nature environment characters from an image precisely. To solve that problem the blind IQEMs are utilized in this paper. To control the computation complexity, only three IQEMs are used here. In future other metrics can be considered.

In this paper, the deep learning-based detection method is employed. Comparing with other detection method, the proposed method has three advantages at least. First, the environment perception-based image enhancement algorithm can select and generate the image with better color, abundant texture, and elaborated edge for the following deep learning-based classification. Second, the computation complexity of proposed algorithm is low which can guarantee the real-time application of multi-rotor UAV. No complex optimization processing is utilized in our proposed method thus the calculation of SSD-based method can achieve a fast recognition speed. Currently, the average processing speed of total algorithm is large than 14-15 frames per second. Third, the detection rate is high [37]. The deep learning-based method has been proved to be one of best detection algorithms in recent years for different applications. In future, other target detection functions, such as the detection mission of the power line or the train track can be developed in this system.

5. Conclusion

An image detection method of manhole cover from the nature scene is proposed in this paper. In contrast to the traditional application, the manhole cover image is captured by a multi-rotor UAV system which means its imaging quality is affected by the nature environment light and shadow seriously. To conquer that problem, first the blind IQEMs are used to filter the improper image data. Then a kind of adaptive MSR method is used to improve its imaging quality. Second, a deep learning tool is used to detect the manhole cover from the ground image. A fast processing method, i.e. the SSD method, is utilized. Third, the spatial coordinates of manhole cover are also estimated from the captured aerial image. Many practical applications have shown the validity of proposed system and method.

References

- [1] V. Ghase, A. A. Pouyan, and M. Sharifi, "Human activity recognition in smart homes based on a difference of convex programming problem," *KSII Transactions on Internet and Information Systems*, vol. 11, no. 1, pp. 321-344, January, 2017. [Article \(CrossRef Link\)](#)
- [2] G. Abdelbacet, Z. Ghada, S. Mounir, and K. Abdennaceur, "A real time environmental monitoring for smart city surveillance based GUI on Android platform," in *Proc. of IEEE International Multi-Conference on Systems, Signals & Devices*, pp. 1-6, March 16-19, 2015. [Article \(CrossRef Link\)](#)
- [3] L. Yin, C. Liu, X. Lu, J. Chen, and C. Liu, "Efficient compression algorithm with limited resource for continuous surveillance," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 11, pp. 5476-5496, November, 2016. [Article \(CrossRef Link\)](#)
- [4] Y. Yu, J. Li, H. Guan, C. Wang, and J. Yu, "Automated detection of road manhole and sewer well covers from mobile LiDAR point clouds," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 9, pp. 1549-1553, September, 2014. [Article \(CrossRef Link\)](#)
- [5] H. Wang, N. Huo, J. Li, K. Wang, and Z. Wang, "A road quality detection method based on the Mahalanobis-Taguchi system," *IEEE Access*, vol. 6, pp. 29078-29087, May, 2018. [Article \(CrossRef Link\)](#)
- [6] X. Chen, J. Xu, and W. Guo, "The research about video surveillance platform based on cloud computing," in *Proc. of International Conference on Machine Learning and Cybernetics*, pp. 979-983, July 14-17, 2013. [Article \(CrossRef Link\)](#)
- [7] N. D. Hoang, "Detection of surface crack in building structures using image processing technique with an improve Otsu method for image thresholding," *Advanced in Civil Engineering*, vol. 2018, pp. 3924120-1 – 3924120-10, April, 2018. [Article \(CrossRef Link\)](#)
- [8] H. Cui, J. Liu, and G. Su, "Combined static and dynamic platform calibration for an aerial multi-camera system," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 6, pp. 2689-2708, June, 2016. [Article \(CrossRef Link\)](#)
- [9] Y. Yu, H. Guan, and Z. Ji, "Automated detection of urban road manhole covers using mobile laser scanning data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3258-3269, December, 2015. [Article \(CrossRef Link\)](#)
- [10] G. Jia, G. Han, H. Rao, and L. Shu, "Edge computing-based intelligent manhole cover management system for smart cities," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1648-1656, June, 2018. [Article \(CrossRef Link\)](#)
- [11] A. Giyenko, and Y. I. Cho, "Intelligent UAV in smart cities using IoT," in *Proc. of International Conference on Control, Automation and Systems*, pp. 207-210, October 16-19, 2016. [Article \(CrossRef Link\)](#)
- [12] G. Zhang, L. Wang, Z. Zheng, Y. Chen, Z. Zhou, and K. Zhao, "No-reference aerial image quality assessment based on natural scene statistics and color correlation blur metric," in *Proc. of IEEE PES International Conference on Transmission & Distribution Construction, Operation & Live-Line Maintenance*, pp. 1-4, September 12-15, 2016. [Article \(CrossRef Link\)](#)
- [13] X. Zhao, Q. Fei, and Q. Geng, "Vision based ground target tracking for rotor UAV," in *Proc. of IEEE International Conference on Control and Automation*, pp. 1907-1911, June 12-14, 2013. [Article \(CrossRef Link\)](#)
- [14] C. Kim, D. S. Han, J. K. Kim, and B. I. Kim, "Automatic detection of defective welding electrode tips using color segmentation and Hough circle detection," in *Proc. of IEEE Region 10 Conference*, pp. 1371-1374, November 22-25, 2016. [Article \(CrossRef Link\)](#)
- [15] H. Liu, H. Lu, and Y. Zhang, "Image enhancement for outdoor long-range surveillance using IQ-learning multiscale Retinex," *IET Image Processing*, vol. 11, no. 9, pp. 786-795, September, 2017. [Article \(CrossRef Link\)](#)
- [16] L. Wang, X. Yao, Z. Meng, T. Liu, Z. Li, B. Shi, Y. Su, R. Zhang, and W. Liu, "An optical coherence tomography attenuation compensation algorithm based on adaptive multi-scale Retinex," *Chinese Journal of Lasers*, vol. 40, no. 12, pp. 1204001-1 – 1204001-6, December, 2013.

- [17] A. Carrio, C. Sampedro, A. Rodriguez-Ramos, and P. Campoy, "A review of deep learning methods and applications for unmanned aerial vehicles," *Journal of Sensors*, vol. 2017, pp. 3296874-1 – 3296874-13, August, 2017. [Article \(CrossRef Link\)](#)
- [18] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single Shot MultiBox Detector," in *Proc. of European Conference on Computer Vision*, pp. 21-37, December, 2016. [Article \(CrossRef Link\)](#)
- [19] Z. Wu, L. Huang, D. Hu, and C. Ding, "Ground resolution analysis based on gradient method in geosynchronous SAR," in *Proc. of IEEE International Conference on Signal Processing, Communication and Computing*, pp. 1-4, August 5-8, 2013. [Article \(CrossRef Link\)](#)
- [20] H. Liu, W. Wang, F. Gao, Z. Liu, Y. Sun, and Z. Liu, "Development of space photographic robotic arm based on binocular vision servo," in *Proc. of International Conference on Advanced Computational Intelligence*, pp. 345-349, October 19-21, 2013. [Article \(CrossRef Link\)](#)
- [21] C. Ramirez-Atencia, V. Rodriguez-Fernandez, A. Gonzalez-Pardo, D. Camacho, "New artificial intelligence approaches for future UAV ground control stations," in *Proc. of IEEE Congress on Evolutionary Computation*, pp. 2775-2782, June 5-8, 2017. [Article \(CrossRef Link\)](#)
- [22] T. Jiang, J. Li, B. Li, K. Huang, C. Yang, and Y. Jiang, "Trajectory optimization for a cruising unmanned aerial vehicle attacking a target at back slope while subjected to a wind gradient," *Mathematical Problems in Engineering*, vol. 2015, pp. 635395-1 – 635395-14, June, 2015. [Article \(CrossRef Link\)](#)
- [23] Z. Zhang, and K. Li, "Study on algorithm for panoramic image basing on high sensitivity and high resolution panoramic surveillance camera," in *Proc. of IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 359-364, August 27-30, 2013. [Article \(CrossRef Link\)](#)
- [24] A. K. Moorthy, and A. C. Bovik, "Blind image quality assessment: from natural scene statistics to perceptual quality," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3350-3364, December, 2011. [Article \(CrossRef Link\)](#)
- [25] H. Liu, W. Wang, Z. He, Q. Tong, X. Wang, W. Yu, and M. Lv, "Blind image quality evaluation metrics design for UAV photographic application," in *Proc. of IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems*, pp. 293-297, June 8-12, 2015. [Article \(CrossRef Link\)](#)
- [26] S. Jaiswal, and M. Valstar, "Deep learning the dynamic appearance and shape of facial action units," in *Proc. of IEEE Winter Conference on Applications of Computer Vision*, pp. 1-8, March 7-10, 2016. [Article \(CrossRef Link\)](#)
- [27] S. Bianco, M. Buzzelli, D. Mzaaini, and R. Schettini, "Deep learning for logo recognition," *Neurocomputing*, vol. 245, pp. 23-30, July, 2017. [Article \(CrossRef Link\)](#)
- [28] J. Chang, H. Jiang, Z. Weng, X. Cong, and Y. Jin, "Design of wide angle space optical systems of long focal length," *Acta Armamentarii*, vol. 24, no. 1, pp. 42-44, January, 2003.
- [29] H. Liu, C. Wang, H. Lu, and W. Yang, "Outdoor camera calibration method for a GPS & PTZ camera based surveillance system," in *Proc. of IEEE International Conference on Industrial Technology*, pp.263-267, March 14-17, 2010. [Article \(CrossRef Link\)](#)
- [30] D. Mery, E. Svec, M. Arias, V. Rizzo, J. M. Saavedra, and S. Banerjee, "Modern computer vision techniques for X-ray testing in baggage inspection," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 4, pp. 682-692, April, 2017. [Article \(CrossRef Link\)](#)
- [31] Y. Wang, and J. Zheng, "Real-time face detection based on YOLO," in *Proc. of IEEE International Conference on Knowledge Innovation and Invention*, pp. 221-224, July 23-27, 2018. [Article \(CrossRef Link\)](#)
- [32] P. Dong, and W. Wang, "Better region proposals for pedestrian detection with R-CNN," in *Proc. of Visual Communications and Image Processing*, pp. 1-4, November 27-30, 2016. [Article \(CrossRef Link\)](#)
- [33] R. U. Khan, X. Zhang, R. Kumar, and H. A. Tariq, "Analysis of resnet model for malicious code detection," in *Proc. of International Computer Conference on Wavelet Active Media Technology and Information Processing*, pp. 239-242, December 15-17, 2017. [Article \(CrossRef Link\)](#)

- [34] K. Shi, H. Bao, and N. Ma, "Forward vehicle detection based on incremental learning and fast R-CNN," in *Proc. of International Conference on Computational Intelligence and Security*, pp. 73-76, December 15-18, 2017. [Article \(CrossRef Link\)](#)
- [35] B. Liu, W. Zhao, and Q. Sun, "Study of object detection based on faster R-CNN," in *Proc. of Chinese Automation Congress*, pp. 6233-6236, October 20-22, 2017. [Article \(CrossRef Link\)](#)
- [36] F. Ye, C. Chen, Y. Lai, and J. Chen, "Fast circle detection algorithm using sequenced Hough transform," *Optics and Precision Engineering*, vol. 22, no. 4, pp. 1104-1111, April, 2014. [Article \(CrossRef Link\)](#)
- [37] W. Sultani, S. Mokhtari, and H.-B. Yun, "Automatic pavement object detection using superpixel segmentation combined with conditional random field," *IEEE Transactions on Intelligent Transaction Systems*, vol. 19, no. 7, pp. 2076-2085, July, 2018. [Article \(CrossRef Link\)](#)



Haoting Liu received M.S. degree from the department of aerospace ergonomics, Astronaut Research and Training Center of China in 2005, and Ph.D. degree from the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences in 2010. Recently, he is with the School of Automation and Electrical Engineering, University of Science and Technology Beijing as an associate professor. His major research interests include pattern recognition, computer vision, image processing, data mining, physiological signal processing, and ergonomics.



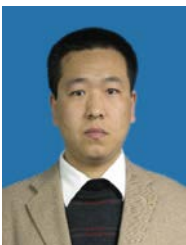
Beibei Yan received M.S. degree from Beijing University of Aeronautics and Astronautics in 2015. She is currently with Beijing Institute of Aerospace Control Device as a junior engineer. Her research interests include image analysis, data mining, and optical engineering.



Wei Wang received M.S. degree from Beijing University of Aeronautics and Astronautics in 1991 and Ph. D. degree from China Academy of Launch Vehicle Technology in 1998. In 2013, he was selected as the academician of the Chinese Academy of Sciences. He is currently with department of R&D, China Aerospace Science and Technology Corporation, as a research fellow. His main research scopes include optoelectronic technique and fiber-optic sensor.



Xin Li received M.S. degree from Aerospace Engineering University of PLA Strategic Support Force in 2009. He is currently with Jiuquan Satellite Launch Center. His research interests include UAV application and data communication.



Zhenhui Guo received M.S. degree from Peking University in 2009. He is currently with Jiuquan Satellite Launch Center. His research interests include image processing and application.