Recent progress in wide-area surveillance: protecting our pipeline infrastructure

Vijayan K. Asari[a], Sidike Paheding[a], Chen Cui[a], and Varun Santhaseelan[b]

[a] University of Dayton, 300 College Park, OH, USA
[b] Auviz Systems, Campbell, CA, USA

ABSTRACT

The pipeline industry has millions of miles of pipes buried along the length and breadth of the country. Since none of the areas through which pipelines run are to be used for other activities, it needs to be monitored so as to know whether the right-of-way (RoW) of the pipeline is encroached upon at any point in time. Rapid advances made in the area of sensor technology have enabled the use of high end video acquisition systems to monitor the RoW of pipelines. The images captured by aerial data acquisition systems are affected by a host of factors that include light sources, camera characteristics, geometric positions and environmental conditions. We present a multistage framework for the analysis of aerial imagery for automatic detection and identification of machinery threats along the pipeline RoW which would be capable of taking into account the constraints that come with aerial imagery such as low resolution, lower frame rate, large variations in illumination, motion blurs, etc. The proposed framework is described from three directions. In the first part of the framework, a method is developed to eliminate regions from imagery that are not considered to be a threat to the pipeline. This method makes use of monogenic phase features into a cascade of pre-trained classifiers to eliminate unwanted regions. The second part of the framework is a part-based object detection model for searching specific targets which are considered as threat objects. The third part of the framework is to assess the severity of the threats to pipelines in terms of computing the geolocation and the temperature information of the threat objects. The proposed scheme is tested on the real-world dataset that were captured along the pipeline RoW.

Keywords: Pipelines, image segmentation, background elimination, part-based, wide-area surveillance

1. INTRODUCTION

Traditionally, pipeline surveillance is conducted qualitatively by aircraft, driving patrol, and walking inspection to record features along the RoW that are important to the pipeline's safety and security. These manual techniques produce results that may not be sensitive or reach desired accuracy to localize or subtle problem identification. Considering the vast amount of area to be monitored in regions with less population, aerial monitoring is found to be the most viable option.

Rapid advances made in the area of camera and sensor technology have enabled the use of video acquisition systems to monitor the RoW of pipelines. Huge amount of data is thus made available for analysis. However, it would be very expensive to employ analysts to scan through the data and identify threats to the RoW in the vast amount of wide area imagery. This warrants the deployment of an automated mechanism that is able to detect threats to the RoW and send out warnings when threats are detected.

Machinery objects, such as construction equipment and heavy vehicles, have been major threats to pipeline infrastructure. Several vehicle detection algorithms have been proposed in the literature. Zhao and Nevatia [1] effectively utilized car body, edges of the front windshield, and the shadow as the features for cat detection. Moon et al. [2] introduces a simple vehicle detection algorithm by exploring four elongated edge operators. A top-down matching method is developed for vehicle detection from high resolution aerial imagery [3]. Grabner et al. [4] exploits on-line boosting with interactive training framework for automatic car detection. Sahli et al. [5] present an alternative approach to the car detection using scale-invariant feature transform features and affinity propagation algorithm. Recently, a three-stage pattern recognition framework is proposed to detect construction equipment in various lighting conditions and different object orientations [6].

However, the majority of these techniques either computationally expensive or suffer from complex environment in aerial imagery, and neither of them consider the potential security issue of the detected objects for wide area surveillance. Thus, we present a multistage framework for the analysis of aerial imagery for automatic detection and identification of machinery threats along the pipeline right of way which would be capable of taking into account the constraints that come with aerial imagery such as low resolution, lower frame rate, large variations in illumination, motion blurs, etc.

The rest of paper is organized as follows. In section 2, an implementation framework of the proposed scheme is provided. In section 3, experimental results are presented and discussed. Finally, section 4 outlines concluding remarks and future research direction in this technology.

2. PROPOSED SCHEME

The proposed machinery threat detection technology can be categorized into three phases, namely background elimination, part-based object detection and risk assessment. Figure 1 depicts flow diagram of the proposed scheme.
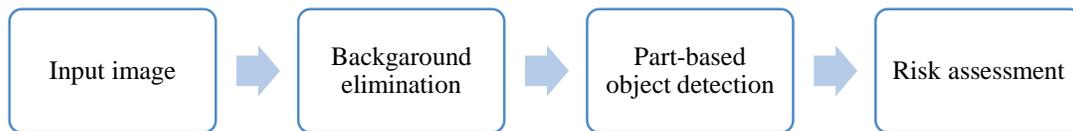

Figure 1: Block diagram of the proposed scheme.

2.1 Background elimination

The aim of developing the background elimination model was to provide information regarding the contents of an image that could be used in the process of threat detection. Some of the key observations made in the study are: (a) aerial imagery consists of various kinds of regions, (b) the regions can be segmented based on the information content in image domain or in a transformed domain. It is observed that plain ground does not contain much information contents, while buildings in an image have strong edge features and the trees have strong textural contents. Based on these observations, an algorithm is designed to efficiently segment regions in an image.

On the other hand, during the process of monitoring pipeline through a small aircraft, the experienced observers will adaptively eliminate most objects from their vision system that are no recognized as threats, such as houses, tress, etc., are less likely be a threat to pipeline. To mimic this kind of human vision system, we propose an automatic background elimination algorithm which can be broken into two parts: local textural features based segmentation (LTFS) and adaptive perception based segmentation (APS). The advantage of developing an automatic background elimination technique can be summarized as follows:

- Eliminate background in aerial imagery for a faster threat identification.
- Extract semantic information from scenes that can aid in threat detection.
- Utilize context cues to identify proper landmarks for better accuracy during change detection processes.
- Gather intelligence from a scene to aid in decision making for users.

2.1.1 LTFS

Image segmentation plays an important role in enhancing the object detection rate. We herein introduce a new segmentation method, named LTFS, which is expected to improve both the accuracy and efficiency of our present threats detection algorithm. The LTFS is based on the property of the neighborhood area around every pixel within an image. The output of the LTFS only contains prominent information of the input image, such as abnormal regions or full connected inhomogeneous area.

The concept of the proposed algorithm is illustrated in Fig. 2. Let P be a point on the edge of an object in an image, and the edge separates the pixel points into two groups, so that the neighbor pixel around P can be separated into two classes. Each class has the same intensity value as shown in Fig. 2 as represented in two colors. Thus, the average intensity of all the neighbor pixels will be larger than the intensity values of one group of pixel points, and less than the other group of pixel

points. If the neighbor pixels are thresholded by the average intensity, one group of the points will be 1, the other group will be 0, so that the sequence of P, contributed by the threshold pixels, will be a uniform pattern [7] in the circular direction.
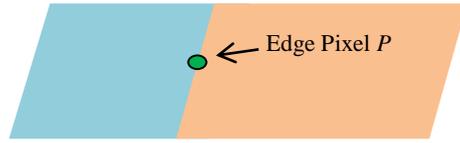

Figure 2: The concept of the LTFS algorithm.

For a given image, let I_p ($p = 1,2, ... ,8$) be the intensity value of a pixel in a 3×3 neighborhood. Then the average intensity of the related neighbor pixels is computed by

$$I_{ave} = \frac{I_1 + I_2 + ... + I_6 + I_7 + I_8}{8}.$$ (1)

If $I_p > I_{ave}$, $I_p = 1$, otherwise $I_p = 0$, expressed as

$$I_p = \begin{cases} 1 & I_p > I_{ave} \\ 0 & otherwise \end{cases}.$$ (2)

Thus, the related neighbor pixel values are 0 and/or 1. Let I_{new_p} be the new value of the pth neighbor pixel (1 or 0), then, the new value of the center pixel is concatenated by all the neighbor pixels, expressed by

$$I_{new_c} = I_{new_1} I_{new_2} I_{new_3} ... I_{new_6} I_{new_7} I_{new_8}.$$ (3)

If I_{new_c} is a uniform pattern (except 00000000 and 11111111), then $I_{new_c} = 1$, otherwise, $I_{new_c} = 0$. The last step of the LTFS is to perform morphological operation to remove imperfections in the binary image.

Even though the threat objects are not single intensity object, the intensity levels of the most pixel points have less differences so that if the difference between the average intensity level and the neighbor pixel's intensity level is within a small range, we consider the neighbor pixels as one intensity level. The output of the LTFS only contains prominent information of the input image, such as abnormal regions or full connected inhomogeneous area. As shown in Fig. 3, the majority of background is eliminated and only some protruding regions are remained which indicates the possible location of the target.
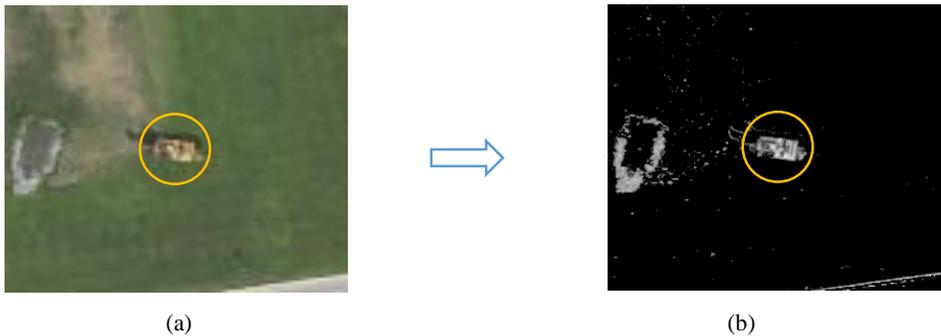

(a)   (b)

Figure 3: Sample output of the LTFS algorithm. (a) Original RGB image, and (b) LTFS output. (Yellow circle: target location)

2.1.2 APS

The LTFS method provides global background elimination. However, it cannot be trained to eliminate specific regions in a given image. Therefore it is necessary to develop an advanced algorithm for semantic segmentation purpose. Thus, we propose the APS which is an artificial neural net based segmentation algorithm that can be trained to segment out specific objects from images.
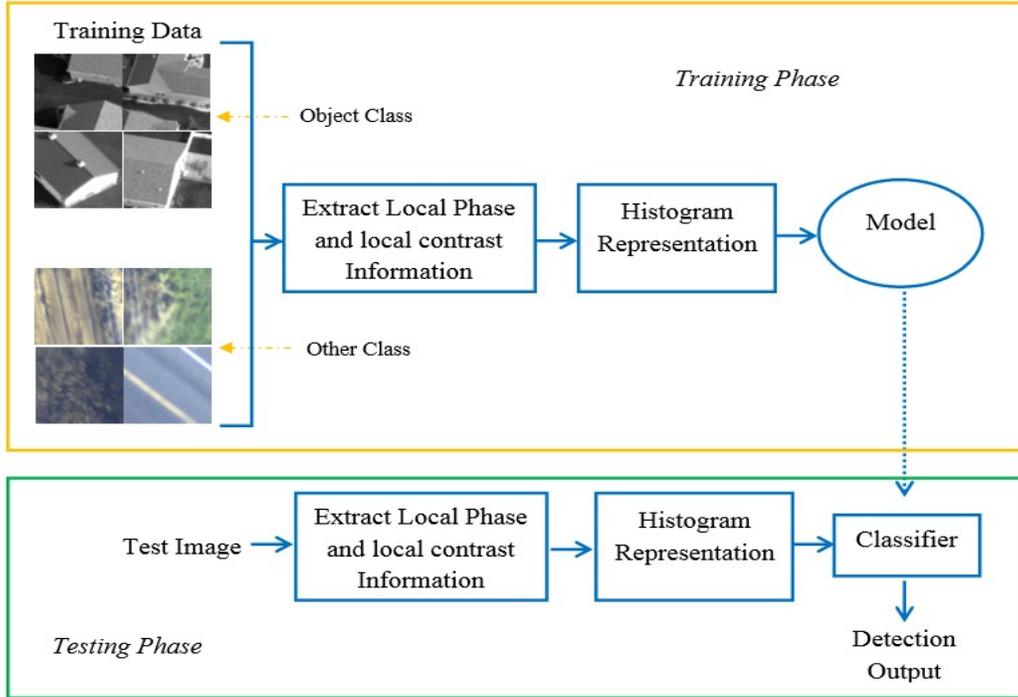

Figure 4: Architecture for the proposed APS algorithm.

The idea of APS model comes from the key observations in aerial imagery. One of the main observations during data analysis is the fact that most of the regions in the image do not contain a lot of information. There are also a lot of regions where the probability of finding threats are considerably low. In order to reduce the computational load on the object detection, we design a framework to segment out non-salient regions of an input image. A complete architecture of the algorithm is shown in Fig. 4. A test image is divided into various segments and passed through the trained model to detect the presence of an object. In Fig. 4, the local phase and local contrast are contextual features that are computed from the monogenic signal [8], expressed separately by

$$A(\boldsymbol{x}) = \sqrt{f^2(\boldsymbol{x}) + f_1^2(\boldsymbol{x}) + f_2^2(\boldsymbol{x})} \tag{4}$$

and

$$\varphi(\boldsymbol{x}) = arctan\left(\frac{\sqrt{f_1^2(\boldsymbol{x}) + f_2^2(\boldsymbol{x})}}{f(\boldsymbol{x})}\right), \varphi \in [0, \pi) \tag{5}$$

where $A(\boldsymbol{x})$ represents the local phase, and $\varphi(\boldsymbol{x})$ is the local amplitude. To obtain $f(\boldsymbol{x})$, $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$, we assume that an image $S(\boldsymbol{x})$ is represented by

$$S(\boldsymbol{x}) = A(\boldsymbol{x})\cos(\varphi) \tag{6}$$

where $x = (x, y)$ is the spatial coordinates of the signal S. Then if S is convolved with the transform function of even and odd pairs of spherical quadrature filters (SQFs) as shown in Eqs. (7), (8) and (9), we can obtain the components of the monogenic signal representation $(f(x), f_1(x), f_2(x))$.

$$f(x) = S(x) * g_e(x) \qquad (7)$$

$$f_1(x) = S(x) * g_{o1}(x) \qquad (8)$$

$$f_2(x) = S(x) * g_{o2}(x) \qquad (9)$$

where '$*$' represents the 2D convolution, $g_e(x)$ is the spatial domain representations of log Gabor filter, and $g_{o1}(x)$ and $g_{o2}(x)$ are the odd set of SQFs, respectively. In terms of physical interpretation, the local phase contains the structure information of the objects while the local contrast information is represented by the local amplitude. In this research, the local phase and the local amplitude are used to represent regions of the image both in training and testing phases. For illustration, a sample result of the APS algorithm is shown in Fig. 5. In this specific example, buildings are being segmented.
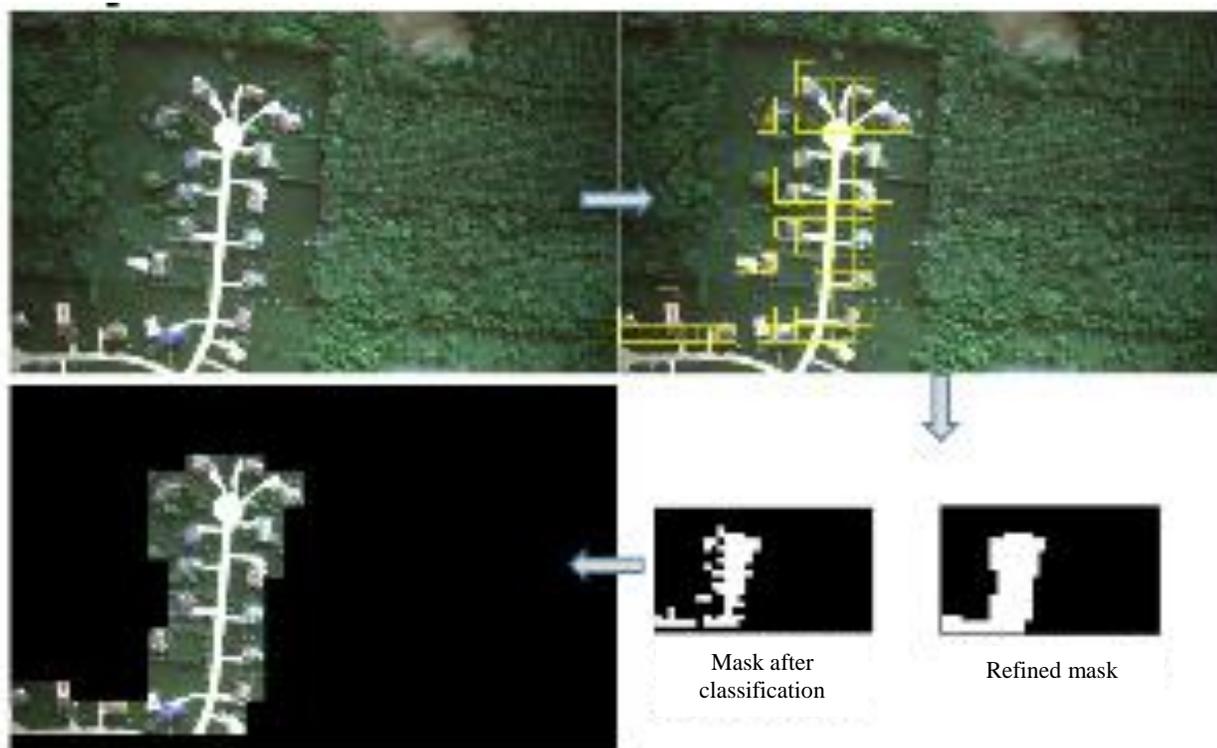

Mask after classification

Refined mask

Figure 5: Illustration of the APS segmentation.

2.2 Part-based object detection

In aerial imagery, a major challenge for detection is when the object of interest is partially occluded by shrubs, trees, buildings, etc. The part-based model has been shown attractive performance in object recognition due to its ability to cope with partial occlusions and large appearance variations [9-10]. Our proposed part-based model is demonstrated in Fig. 6. At first, an object is partitioned into a certain number of parts that varying by the size of object, then local phase information is used to extract informative attributes for describing individual parts. Next, object parts represented using local phase are converted into a larger number of clusters, similar parts are grouped into same cluster and then represented by histogram of oriented phase to describe specific pattern of the parts. The next step is to organize each of the detected parts and their

attributes as an integrated entity. Since a target can be represented by certain number of patterns, we can train a classifier to detect such local patterns of the target, so that an occluded object can be detected by parts in the input scene.
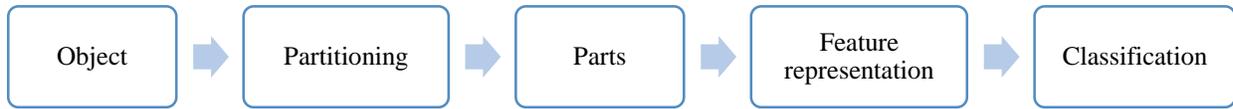

Figure 6: Flowchart of the proposed part-based model.

2.3 Risk assessment

The part- based object detection technique outputs the pixel location of the threat object in the input image. However, in real world, a pipeline operator has to know the geolocation of an object for preventing any damage to the pipeline. This requires a registration process between images and geographical map. In addition, some detected machinery threats may be placed far away from the pipeline or even they are not being operated where the probability of that to be a threat is significantly low. Considering this issue, we designed a framework, as shown in Fig. 7, which can automatically analyze the geolocation and temperature of the detected object and assign risk level of a threat as high, medium and low.
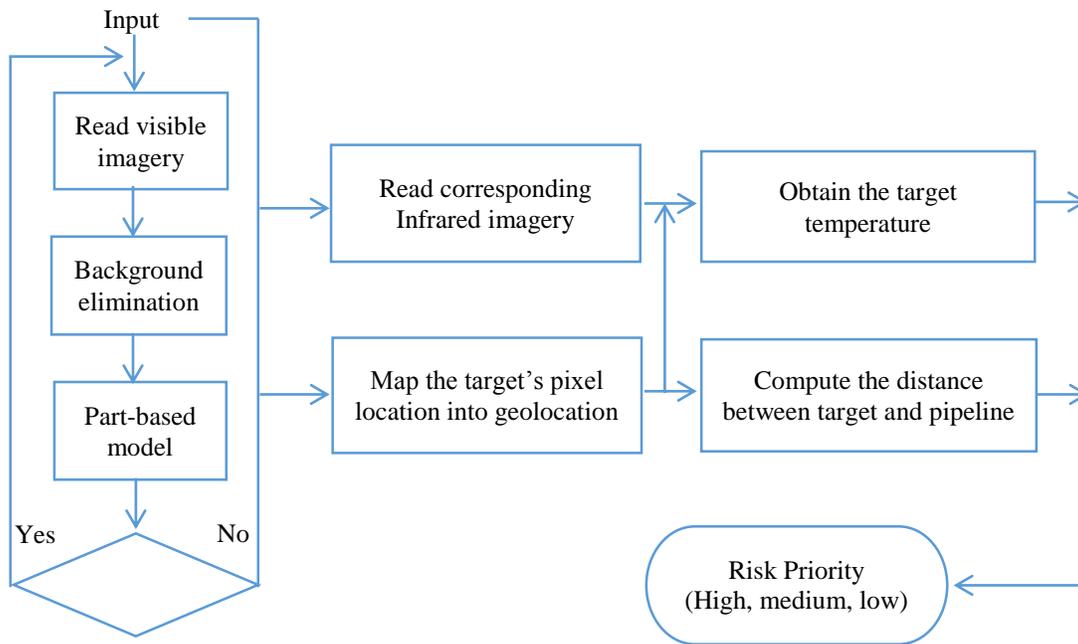

Figure 7: The proposed framework for risk assessment.

If a threat is assigned as a "high" which delivers a meaning that this object is the more potential threat to the pipeline, whereas if it is marked as a "low", then the detected object has less risk to the pipeline. In Fig. 8, assume that a detected object located in P_1 in an input image (rectangular region in blur color), and P_2 is the nearest point to P_1 and it locates on the pipeline centerline with geo coordinates, then we compute the shortest distance between pipeline centerline and the object, denoted as D. Notice that the coordinates of P_1 is the spatial location in the image. In order to compute the physical distance between P_1 and P_2, we need to convert the pixel coordinates of P_1 into geo coordinates. Since we know the geo coordinates of pixels in the four-corner of the image, we can easily find a transformation matrix to map image spatial location to geolocation, so that the geo coordinates of the object will be attained. Moreover, the temperature information of the target is obtained using the pixel value of the object in corresponding infrared imagery.
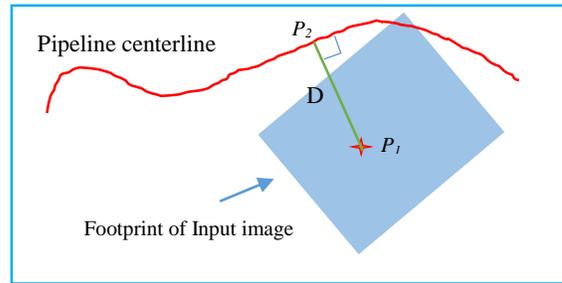
Figure 8: Demonstration of the distance calculation method.

To evaluate accuracy of the proposed distance measurement technique, we embedded five sample targets in testing images, and the distances from targets to the pipeline are provided by Global Mapper, which is used as a ground truth for our analysis. The comparison of our proposed method with the ground truth is shown in Table 1.

Table 1: Distance calculation statistic.

Method / Object	Ground truth (feet)	Proposed (feet)	Average Mean Square Error
Target 1	7.897684	7.928991	
Target 2	75.533661	75.173040	
Target 3	95.846070	95.086690	0.1768 feet
Target 4	110.473560	110.194510	
Target 5	62.576181	62.262255	

3. EXPERIMENTAL RESULTS

In this section, we will show results of our automatic machinery threat detection technique on a real-world dataset. The images in the database have been captured at altitude around 1000 feet along the pipeline RoW and data capture was done by one infrared and one visible cameras pointing towards the pipeline centerline. The objective in this research is develop a full-fledged system that can automatically detect potential machinery threats and aid human analysts for subsequent actions. The results of our proposed method are presented in two stages. The first stage shows the performance of the proposed background elimination technique in varying background conditions. The second stage presents the detection output using the proposed part-based model after background elimination. Risk assessment results will be generated as a text file after these two stages, however, we are not showing it here.

3.1 Results of background elimination

In Fig. 9, the results are shown in sequential order for the proposed LTFS and APS algorithms. Fig. 9 (a) is a sample test image which contains a threat object (red circled) that closes to the pipeline RoW. Fig. 9 (b) shows the output of the LTFS algorithm, as seen in the result, most of undesired background has been removed but the object is kept in the output image. Next, the APS is applied to the output image of the LTFS. Fig. 9 (c) shows the local phase image which was used for training and testing during APS process as mentioned in Section 2.1.2. Fig. 9 (d) is the final output after LTFS and APS processes. At seen in Fig. 9 (d), there are only few regions of the original image are left, this output would significantly contribute to the object detection stage since only few patches of the image will be considered for searching the object.

3.2 Part-based object detection

After background elimination, the part-based object detection model described in Section 2.2 is used for threat object detection. In this model, the sliding windows technique is used for scanning the image while SVM is employed for finding the object. During the sliding windows, due to the background segmentation technique eliminates most of non-target regions
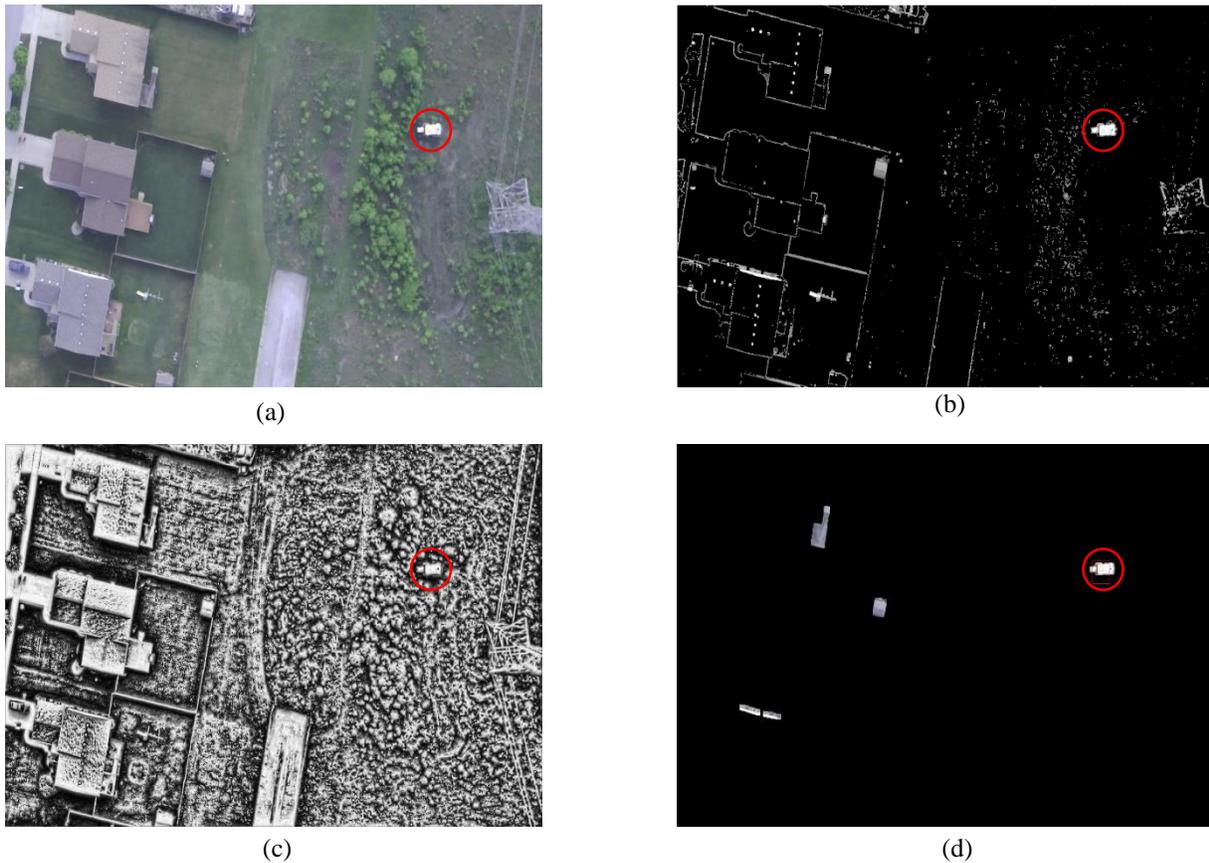
(a)

(b)

(c)

(d)

Figure 9: Results of background elimination. (a) Original image, (b) LTFS, (c) local phase, and (d) LTFS+APS.

and sets the intensities of those regions are zero as shown in Fig. 10 (b), only if the amount of non-zero intensity values great that 30% in a local region is computed, this accelerates the processing speed as well as the detection rate. Figure 10 shows a sample result using the proposed part-based technique. As shown in Fig. 10 (c), the multiple parts of the object are detected without any false alarm.

4. CONCLUSIONS

In this paper, we have presented a new automated monitoring system to mimic human vision system on the application of machinery threat detection on pipeline RoW. The proposed technique has been simulated and tested on real-life dataset under various challenging conditions to investigate its reliability and feasibility. The proposed system yields above 85% accuracy for machinery threat detection and offers a practical candidate for a wide area surveillance to protect our pipeline infrastructure. Currently, work is in process to refine algorithm with respect to computation speed as well as accuracy. In addition, we are establishing a standard databased for construction vehicles detection which will be available to public soon.
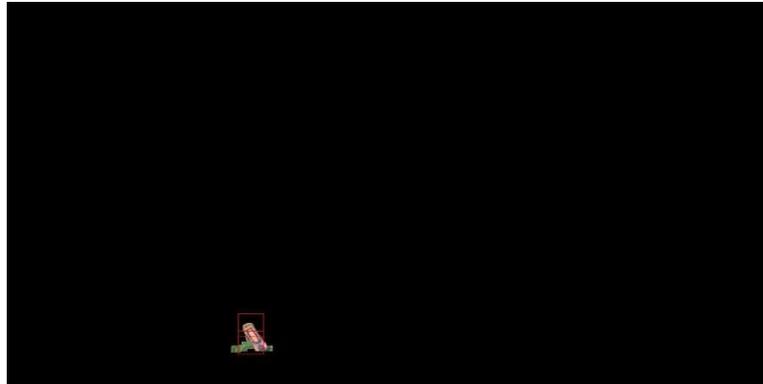
ACKNOWLEDGMENTS

(a)

(b)

(c)

Figure 10: The detection output of the proposed algorithm. (a) Original image (yellow circle: object location), (b) after background elimination, and (c) part-based detection output (red rectangular: multiple parts are detected).

5. REFERENCES

[1] Zhao, T. and Nevatia, R., "Car detection in low resolution aerial image," Proc. Eighth IEEE International Conference on Computer Vision (ICCV), 710-717 (2001).

[2] Moon, H., Chellappa, R. and Rosenfeld, A., "Performance analysis of a simple vehicle detection algorithm," Image and Vision Computing, 20(1), 1-13 (2002).

[3] Hinz, S., "Detection of Vehicles and Vehicle Queues in High Resolution Aerial Images," In Photogrammetrie - Fernerkunding - Geoinformation (PFG) 3, 201–213 (2004).

[4] Grabner, H., Nguyen, T. T., Gruber, B., Bischof, H., "On-line boosting-based car detection from aerial images," ISPRS Journal of Photogrammetry and Remote Sensing, 63(3), 382-396 (2008).

[5] Sahli, S., Ouyang, Y., Sheng, Y., Lavigne, D. A., "Robust vehicle detection in low-resolution aerial imagery," Proc. SPIE 7668, Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications VII, 76680G (2010).

[6] Nair, B. M., Santhaseelan, V., Cui C. and Asari, V. K., "Intrusion detection on oil pipeline right of way using monogenic signal representation," Proc. SPIE 8745, Signal Processing, Sensor Fusion, and Target Recognition XXII, 87451U (2013).

[7] Topi, M., Timo, O., Matti, P., Maricor, S., "Robust texture classification by subsets of local binary patterns," Proc. of 15th International Conference on Pattern Recognition, 935-938 (2000).

[8] Felsberg, M. and Sommer, G., "The monogenic signal," IEEE Transactions on Signal Processing, 49(12), 3136–3144 (2001).

[9] Felzenszwalb, P. F., Girshick, R. B., McAllester D., and Ramanan, D., "Object detection with discriminatively trained part based models," IEEE Trans. Pattern Anal. Mach. Intell. 32(9), 1627–1645 (2010).

[10] Felzenszwalb, P. F., Girshick R. B. and McAllester, D., "Cascade object detection with deformable part models," IEEE international conference in computer vison and pattern recognition, 2241–2248 (2010).