Spatiotemporal segmentation of moving-object using evaluation of boundary in infrared videos

Lu Liubing
CNGC North Laser Technology Group Co., LTD
Yangzhou, China
e-mail: luliubing@sohu.com

Min Chaobo
School of Electronic and Optical Engineering
Nanjing University of Science and Technology
Nanjing, China
minchaobo@163.com

Xu Hui
School of Electronic and Optical Engineering
Nanjing University of Science and Technology
Nanjing, China
e-mail: wolf7463@sina.com

He Ye
School of Electronic and Optical Engineering
Nanjing University of Science and Technology
Nanjing, China
784361613@qq.com

Abstract—A new method is presented for spatiotemporal segmentation of moving-object using evaluation of boundary in infrared video. At first, the ideal seeds of every moving object are extracted based on the “holes” effect of temporal difference, respectively. The focus of our wok is spatial segmentation. On the basis of the relationship between the global and local standard deviation of seeds, the segmented masks can be grown form the ideal seeds by using different growing thresholds. For determination of the best growing threshold, a criterion is constructed for evaluating the boundary of the segmented mask of infrared moving-object without prior knowledge. According to the proposed criterion, an iterative model which is *segmentation-evaluation-segmentation-evaluation* and the search method called as *coarse to fine* are applied to find the best growing threshold. Meanwhile the best segmented mask is obtained too. The experiment results show that the proposed method is superior and effective on segmentation of moving object in infrared video.

Keywords- ***image segmentation; performance evaluation; infrared video; seed extraction; regional growth***

I. INTRODUCTION

Moving target detection and segmentation is an important part of computer vision. The infrared thermal imager can identify hidden targets through the smoke because of the sensitivity of the hot target so that the moving target segmentation whose common methods are frame difference method, background difference method and optical flow method that widely used for infrared image. Although frame differential method has good character of real-time, it can't detect complete moving target mask. Optical flow method is quite good but needs large amount of calculation and has to satisfy the hypothesis that moving object surface brightness is constant.

Background subtraction based on Gaussian mixture models(MOG) is now recognized as one of the best algorithm. But MOG method can not sacrifice our need because of the complicated background and infrared image’s low SNR. Literature[5] presents a moving object segmentation method based on Markov random field(MRF)， which use the previous frame (MRF) model to reestimate the current frame, and then use the estimation frame to do segmentation based on spatio-temporal domain and also use the previous frame to compensate and correct the segmentation result. But this method can't split a complete moving object mask, and noise interference is more serious under the condition of low SNR when the target motion amplitude is too large.

In order to decrease the disturbance of infrared image noise towards moving object segmentation and extract complete moving object contour ,this paper proposes a segmentation method of moving object using boundary evaluation , builds the evaluation criterion based on the contrast of infrared target segmentation mask boundary and at last the best segmentation mask is got through an iterative model which is segmentation-evaluation-segmentation-evaluation Finally, The experiment results show that the proposed method is superior and effective on segmentation of moving object in infrared video.

II. EVALUATION CRITERION BASED ON THE CONTRAST OF INFRARED TARGET SEGMENTATED MASK BOUNDARY

For infrared target segmentation mask, the accuracy of the boundary can be used to measure the performance of segmentation mask. Because the boundary reflects how the segmentation mask contour matches the actual moving object contour. We use $I_{original}$ for infrared original image while segmentation mask is defined as I_{mask}. 1 is on behalf of the target area and 0 represents the background region. Firstly, 3×3 Sobel operator was used to detect the boundary of infrared target segmented mask for getting its binarization image. Secondly, boundary point was taken as feature by interlaced scanning. $L=\{\mathbf{z}_k:\mathbf{z}_k\in R^2\}_{k=1}^N$ is used as a feature point set of segmentation mask boundary where N represents the number of them shown in Fig .1 (a), (b).

Separate boundary feature points can't accurately describe performance of the boundary at a particular point. Therefore, for each boundary feature points $\mathbf{z}_k$, a feature circular area V_k was got whose center is $\mathbf{z}_k$ and radius is r . Meanwhile, feature area is divided into the internal and external characteristics region depends on segmentation mask as reference object as well as local image gray feature of internal and external region is defined separately .

$$G(V_k^{in}) = \frac{1}{N_k^{in}} \sum_{q \in V_k} \xi_k^{in}(\mathbf{q}) \cdot I_{original}(\mathbf{z}_k + \mathbf{q}) \quad (1)$$

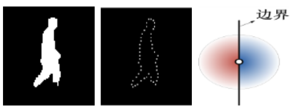
$$G(V_k^{out}) = \frac{1}{N_k^{out}} \sum_{q \in V_k} \xi_k^{out}(\mathbf{q}) \cdot I_{original}(\mathbf{z}_k + \mathbf{q}) \quad (2)$$

In this equation above, $\mathbf{q}$ means the radial vector of feature region V_k , $\xi_k^{in}(\mathbf{q})$ and $\xi_k^{out}(\mathbf{q})$ respectively represent each point's weighting function in internal and external feature area.

$$\xi_k^{in}(\mathbf{q}) = 1[\mathbf{z}_k + \mathbf{q} \in V_k^{in}] \cdot \exp\{-\frac{\|\mathbf{q}\|^2}{2\sigma^2}\} \quad (3)$$

$$\xi_k^{out}(\mathbf{q}) = 1[\mathbf{z}_k + \mathbf{q} \in V_k^{out}] \cdot \exp\{-\frac{\|\mathbf{q}\|^2}{2\sigma^2}\} \quad (4)$$

If $\mathbf{z}_k + \mathbf{q}$ is within V_k^{in} / V_k^{out} , $1[\mathbf{z}_k + \mathbf{q} \in V_k^{x}]$ equals 1 rather than 0.This function shows the contribution of each point on the corresponding area towards its Regional grayscale characteristics Fig .1(c) is an example in feature area.



(a) (b) (c)

Figure 1. An example of feature points of boundary: (a) Segmented mask; (b) Feature points of boundary; (c) Feature region (blue: inner region; red: outer region; the darker the pixel value, the higher the weight).

Next, we can define the segmentation mask boundary's average contrast $C(I_{mask})$:

$$C(I_{mask}) = \frac{1}{N} \sum_{k=1}^{N} \left| \frac{G(V_k^{in})}{G(V_k^{out})} - 1 \right| \quad (5)$$

Target's gray consistency in infrared image is good. Therefore, if the segmentation mask are in complete accord with the actual target area, its boundaries of internal and external characteristics of gray should be the biggest; If there have been split or incomplete segmentation, internal and external gray similarity of segmented mask boundaries will increase. So equation (5) can become:

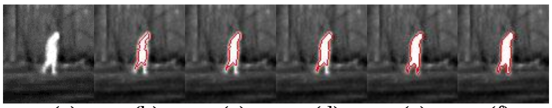
$$\hat{I}_{mask} = \arg \max_{I_{mask}} C(I_{mask}) \quad (6)$$

$\hat{I}_{mask}$ contains the best boundary of segmented mask.

This paper focuses on the segmented mask's own performance rather than whether segmented algorithms are good or not. Therefore, we use the manual selecting method to get seeds and segmented mask with different performance by regional growth in different growth threshold. A comparison between the proposed criterion and misclassification Error(ME) is made to test effectiveness of the boundary evaluation criterion, ME is defined as below:

$$ME = 1 - \frac{|B_o \cap B_r| + |F_o \cap F_r|}{|B_o| + |F_o|} \quad (7)$$

Among this equation, B_o and F_o respectively represents the actual background and the actual target while B_r and F_r mean segmented background and segmented target. ME is a type of segmented evaluation criterion that needs priori knowledge, which shows the false segmented proportion of image pixel. Obviously, the smaller the ME is, the better the segmentation effect we can get.



(a) (b) (c) (d) (e) (f)

Figure 2. Dataset #1: (a) Original image; (b)~(f) Red: the boundaries of the different segmented masks.

TABLE I. THE EVALUATION RESULTS OF DATASET #1

	(b)	(c)	(d)	(e)	(f)
Boundary	0.8789	1.2092	1.2974	1.5409	**1.6977**
ME(%)	1.8137	1.3108	1.1126	0.3353	**0.1067**

(a) (b) (c) (d) (e) (f)

Figure 3. Dataset 2: (a) Original image; (b)~(f) Red: the boundaries of the different segmented masks.

TABLE II. THE EVALUATION RESULTS OF DATASET #2

	(b)	(c)	(d)	(e)	(f)
Boundary	0.885 1	1.121 4	**1.603** **4**	1.371 1	0.5689
ME(%)	3.856 1	2.362 4	**1.082** **2**	1.646 1	48.437 7

ME measures the objective and actual performance of segmented mask .the result of boundaries evaluation criterion from table 1 and table 2 matches well with the subjective perception and objective reality and proves the validity of the boundary evaluation criterion. It's biggest superiority is that prior knowledge is not needed for reference during computing process, thus it can be used for online real-time evalution of segmented mask.

III. SPACE PARTITION BASED ON THE BOUNDARY EVALUATIONS

A. *Seed extraction based on the* ***hole*** *effect*

The frame whose number is k in infrared video is defined as f_k, And the differential figure d_k in time domain segmentation can be explained as:

$$d_k = \left| f_k - f_{k-1} \right| \quad (8)$$

We use the method called Otsu to detect the effective change between pixels, formatting the change detection mask (CDM) where 1 means changing pixels and 0 is for unchanged pixels. The moving target detected by CDM will have *holes* which is the disadvantages of moving target detection in time domain. But the pixel contained in "holes" is exactly the "seed" related to the moving target for the region growing segmentation method. So we will extract these pixels out used as the most meaningful seed point in regional growing most significant points. The specific method is divided into the following steps:

(1)A neighborhood model $[(2\cdot r_a+1)^2-1]$ is used to detect the expanding connected area for CDM. One connected area corresponds to one moving target's changing pixel area where M is the number of moving targets in image: $CDM_i, i=1,2,\ldots\ldots,M$

(2)In CDM_i whether there are two discontinuous changing pixels in each line or each column or not will be tested. If it is, these two points will be connected with straight line;

(3)After all the tests, the point noted to be 0 and intersecting the connecting line is the CDM_i seed for corresponding moving target. Every mask of seed point is defined as S_i .1 is on behalf of the seed point, 0 represents the others.

(4) Repeat steps (2) and (3) till finish extracting seeds from all the moving targets. Here is an example:

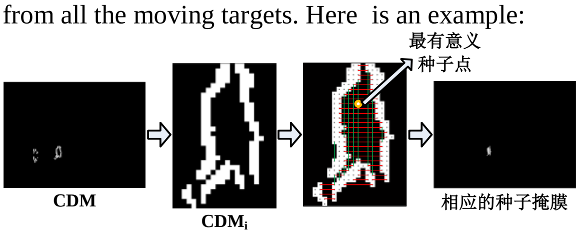


Figure 4. The processing of seed extraction

B. *Regional growth.*

To avoid extra computation and error, we set the centroid marked as1 of pixel region as center in each S_i of moving target. A window of $(2\cdot l_r+1)\times(2\cdot l_c+1)$ is opened where l_r and l_r respectively represent the maximum distance in row direction and column direction of seed's point coordinates in S_i . Window seed mask is Sw_i . Meanwhile, a window called f_k^i is opened in the same position and same size on f_k . The concrete realization method to grow the seeds of a moving object is shown in algorithm 1.T means growth threshold.

Algorithm 1

do

$Q = Sw_i$;

for $x = 2:2\cdot l_r$

for $y = 2:2\cdot l_c$

if $Sw_i(x,y)\neq 1$

A is 8 neighborhood matrix of Sw_i whose center is $Sw_i(x,y)$;

B is 8 neighborhood matrix of f_k^i whose center is $f_k^i(x,y)$;

N_A is the number of elements marked as 1 in A;

if $1\leq N_A\leq 7$

for ii = 1 : 3

for jj = 1 : 3

$C = A(ii,jj)\cdot B(ii,jj)$;

$D(ii,jj)=\left| f_k^i(x,y)-C\right|$;

end

end

D_m =min $\{D(ii,jj)\}$;

$mean(f_k^i)$ =average gray value of f_k^i ;

$D_Z = \left| f_k^i(x,y) - mean(f_k^i)\right|$;

$std(B)$ = standard deviation of B;

$std(f_k^i)$ = standard deviation of f_k^i ;

if $D_m \leq T\cdot std(B)$ and $D_z\leq T\cdot std(f_k^i)$

$Q(x,y)=1$;

end

end

end

end

end

$Sw_i = Sw_i \cup Q$;

While the number of elements marked as 1 in Sw_i is on the rise

The result Sw_i received at the end of the cycle is the corresponding segmented mask under threshold T marked as b.

C. *Search method callede coarse to fine*

Growth threshold T is a very important parameter for region growing which decides the accuracy of the final segmented mask. Because the segmented mask boundary criterion we proposed don't need a priori knowledge for reference so that it can be used to determine the threshold T. We use *coarse to fine* iterative searching method to determine the optimal threshold. For the moving object mask window of a seed, concrete steps are shown as below. t is the iteration steps($t = 0,1,2,\cdots\cdots$) and *step* on behalf of the search step length.

1) *Iterative process of coarse adjustment*：

a) *Set original value* $T(0) = 0$, *step* $= \alpha$ ；

b) $T(t)$ *is used in regional growing of seed mask window* Sw_i *(as is shown in section 2.2), and segmented mask is got.*

c) *The boundary evaluation criteria of* $M_i(t)$ *is calculated;*

d) *When* $t < 1$ ， $T(t+1) = T(t) + step$ ， *repeat steps b* 、 *c and d* ； *when* $t \geq 1$ ， *calculate* $\Delta C = C(M_i(t)) - C(M_i(t-1))$ ；

e) *if* $\Delta C \geq 0$ ， $T(t+1) = T(t) + step$ ， *repeat steps b*、*c and d*；*if* $\Delta C < 0$ ， *stop the iterative process of coarse adjustment and get the initial value of threshold* $T^f = T(t-1)$ $(t = 0)$ *during process of fine adjustment*；

2) *Iterative process of fine adjustment*：

a) *Set original value* $T(0) = T^f$ ， *step* $= \beta$ ；

b) *Repeat steps b*、*c and d*；

c) *if* $\Delta C \geq 0$ ， $T(t+1) = T(t) + 1$ ， *repeat steps b* 、 *c and d* ； *if* $\Delta C < 0$ ， *stop the iterative process of coarse adjustment and get the best value of growing threshold* $T^{best} = T(t-1)$ *.Meanwhile the segmented mask of moving object* $M_i(t-1)$ *is got.*

α and β respectively represent the iteration step length in process of coarse and fine adjustment .This iterative method has the advantage of control threshold and avoid invalid threshold's calculation in order to improve efficiency.

It is important to note that the moving object segmented mask is window images so all moving object's segmented mask should be posted on the corresponding position exactly according to the position of seed mask centroid to merge a complete moving object segmented mask.

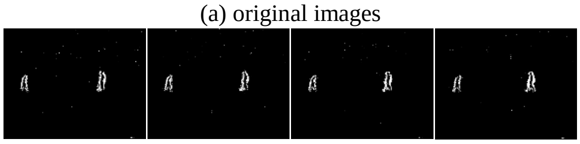
IV. THE EXPERIMENTAL RESULTS AND DISCUSSION

A. *Experimental results*

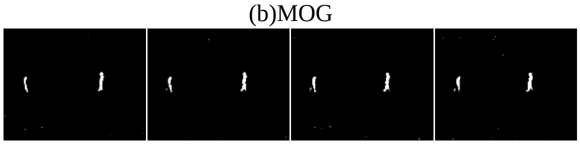
A test of the proposed method is done by using two kinds of infrared video sequence. The No.281 ~ 284 frames in *couple* infrared sequence images is shown in Fig .5 (a), while image size is 240 × 320. The No.91 ~ 94 frames in single of infrared sequence images is shown in Fig .6 (a), and the image size is 400 × 512. In order to verify the effectiveness of this proposed algorithm, a comparison was made between MOG method and moving object segmentation algorithm based on MRF. The experimental results of various kinds of algorithm are shown in Fig .5 and Fig .6.

ME is used to objectively and quantitatively measure performance of various algorithms. The average ME of segmented mask using these three algorithms is respectively shows in Table 3. From the actual comparison of the segmentation results shown in Fig .5 and 6 and the comparison of average ME as we can see in table 3, the algorithm we proposed is superior to other two algorithms no matter in completeness of segmented mask or accuracy of moving targets for both single target and multi-target image.

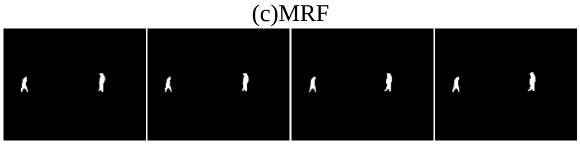
(a) original images



(b)MOG

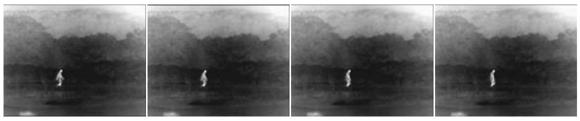


(c)MRF

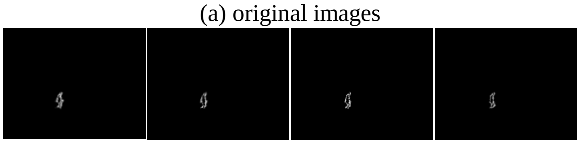


(d) the proposed method

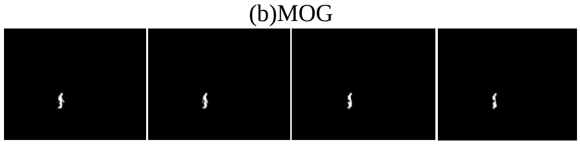
Figure 5. Comparison #1 of experiment results



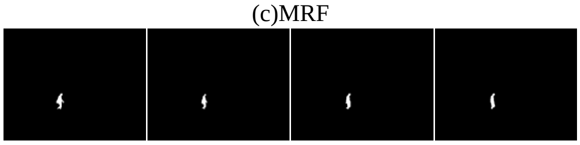
(a) original images



(b)MOG



(c)MRF



(d) the proposed method

Figure 6. Comparison #2 of experiment results

TABLE III. COMPARISON OF AVERAGE ME

	MOG	MRF	the proposed method
double	3.1509%	0.8872%	**0.0328%**
single	2.7163%	0.2076%	**0.0501%**

B. Units analysis and discussion

As is shown in Fig .5(b) and Fig .6(b), the result using MOG is quite accurate. Since it's vulnerable to background interference, the segmented mask is incomplete and has noise interference. Fig .5(c) and 6(c) display the result of MRF method. Image is remarked using MRF first, and then the frame difference is received. The segmentation result is relatively accurate, but segmented mask is still incomplete and has a small amount of noise. It can be seen from Fig .5 (d) and Fig .6 (d) that the method proposed in this paper obtained the most complete segmented mask contour almost without any noise interference. Because we use boundary evaluation and growth way from the inside to the outside so that we can get the exact boundary contour and restrain noise very well.

Specifically, only a set of specific parameters is used in the whole experiment process: radius of feature area neighborhood radius, coarse adjustment step length, fine adjustment step length without any parameter adjustment. The speed of this algorithm dealing with 240 x 320 video image in Core i3-4.01 GHz processor and 2 GB memory's computing platforms reaches 23 frames per second. As for 400 x 512 video image, computing speed reaches 19 frames per second.

V. CONCLUSION

Segmented method of infrared moving object in time domain based on boundary evaluation, using the proposed segmented mask boundary evaluation criterion based on contrast is a good way to overcome the background interference, eliminate noise and accurately extract the moving object boundary contour and effectively reduce the false alarm rate. The method is proved by experiments to be superior and robust. The speed of algorithm must to be improved in the future, and apply this in visible video's moving object detection.

REFERENCES

[1] T. Crivelli, P. Bouthemy, B. Cernuschi-Frías, J. Yao

[2] Simultaneous motion detection and background reconstruction with a conditional mixed-state markov random field

[3] Int J Comput Vision, 94 (3) (2011), pp. 295–316

[4] O. Barnich, M.V. Droogenbroeck

[5] ViBe: a universal background subtraction algorithm for video sequences

[6] IEEE Trans Image Process, 20 (6) (2011), pp. 1709–1724

[7] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[8] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[9] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[10] K. Elissa, "Title of paper if known," unpublished.

[11] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[12] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].