

Target segmentation in non-homogeneous infrared images using a PCA plane and an adaptive Gaussian kernel

Yong Min Kim¹, Ki Tae Park² and Young Shik Moon¹

¹ Hanyang University, Department of Computer Science and Engineering
1271 Sa-3-dong Sangnok-gu, Ansan-si, Kyeonggi-do, Republic of Korea, 426-791
[e-mail: {prime, ysmoon}@hanyang.ac.kr]

² The attached institute of ETRI, National Information Security Academy
P.O. Box 1, Yuseong Daejeon, Republic of Korea, 305-600
[e-mail: kitaep@gmail.com]

*Corresponding author: Young Shik Moon

*Received January 29, 2014; revised May 27, 2015; accepted June 1, 2015;
published June 30, 2015*

Abstract

We propose an efficient method of extracting targets within a region of interest in non-homogeneous infrared images by using a principal component analysis (PCA) plane and adaptive Gaussian kernel. Existing approaches for extracting targets have been limited to using only the intensity values of the pixels in a target region. However, it is difficult to extract the target regions effectively because the intensity values of the target region are mixed with the background intensity values. To overcome this problem, we propose a novel PCA based approach consisting of three steps. In the first step, we apply a PCA technique minimizing the total least-square errors of an IR image. In the second step, we generate a binary image that consists of pixels with higher values than the plane, and then calculate the second derivative of the sum of the square errors (SDSSE). In the final step, an iteration is performed until the convergence criteria is met, including the SDSSE, angle and labeling value. Therefore, a Gaussian kernel is weighted in addition to the PCA plane with the non-removed data from the previous step. Experimental results show that the proposed method achieves better segmentation performance than the existing method.

Keywords: infrared image, segmentation, total least square, principal component analysis, sum of the square error, gaussian weight, error minimization.

1. Introduction

In infrared search and track systems, target segmentation is indispensable in order to exactly track targets in cluttered environments. However, target detection in non-homogeneous backgrounds is a very difficult problem [1], it is necessary to eliminate the non-homogeneous backgrounds in order to improve the performance of the target segmentation. Various methods of target segmentation such as Otsu [2], Normalized cut method [3][4] and fuzzy c-means (FCM) [5][6][7], Saliency method [8] and the Vasquez method have been proposed. Although the Otsu and FCM methods show good performance under a homogeneous background environment, these methods do not effectively distinguish targets from backgrounds under non-homogeneous environments. This is because the brightness values of the non-homogeneous region tend to have irregular values in IR images. Moreover, the brightness value of the non-homogeneous region tends to change gradually from the neighboring pixels. The homogeneous image is largely related to the local information extracted from an image and reflects how uniform a region is [9]. Homogeneous image is defined as the composition of two components [10]: the standard deviation and discontinuity of the intensity within a window. Therefore, homogeneous and non-homogeneous images can be distinguished by using these two components.

F. Dai proposed a practical method for image segmentation by using resonance algorithm to solve the spreading intensity such as gradation types because this algorithm emphasizes the similarity between the adjacent points [11]. To this end, the resonance algorithms of F. Dai tend to have the same or similar features clustered into one region by the spreading of the resonance at points. However, a sudden change of features between the adjacent points can be regarded as the boundary of different regions [12]. Therefore, using the F. Dai's resonance algorithm in homogeneous environments is not appropriate for modeling target segmentation. Vasquez proposed a method of point target detection in homogeneous environments that considers the local mean and the variation values of the pixel values in some windows. Although Vasquez's method can detect a small target in a non-homogeneous environment, it is not suitable for detecting a generic target composed of more pixels. Therefore, by estimating the gradually changed degree, we can estimate the non-homogeneous degree by using the inclined degree from the brightness- plane in 3D space.

As a result, can be solved by estimating how non-homogeneous by using the angle of the principal component analysis (PCA) plane. Therefore, in order to effectively segment the target region from backgrounds under a non-homogeneous environment, we propose a novel method to distinguish the target region exactly given the ROI detected from a non-homogeneous infrared image. The proposed method is composed of three steps. In the first step, we apply a PCA technique by minimizing the total least square (TLS) errors of an IR image. Consequently, the plane vector is generated for estimating the non-homogeneous environment. In the second step, we generate a binary image, consisting of pixels with higher values than the plane. We then calculate the second derivative of the sum of the square errors (SDSSE), an angle between the brightness plane and a plane vector. In the final step, in order to exactly segment the target region, an iteration is performed until the convergence criteria is met. The criteria consist of three conditions: SDSSE, angle and labeling value. Therefore, if do not meet the criteria, the Gaussian weight value is the weighted addition in the PCA plane with the non-removed data from the previous step. Consequently, the data that do not meet the criteria are removed.

This paper is organized as follows. Section 2 describes the proposed method of target segmentation. Section 3 shows the experimental results, and Section 4 includes the concluding remarks.

2. Proposed method

Fig. 1 shows a flow chart of the proposed method for target segmentation with a non-homogeneous environment in the IR image. Since the IR images contain cluttered backgrounds and noises under a non-homogeneous environment, it is difficult to segment the target regions exactly.

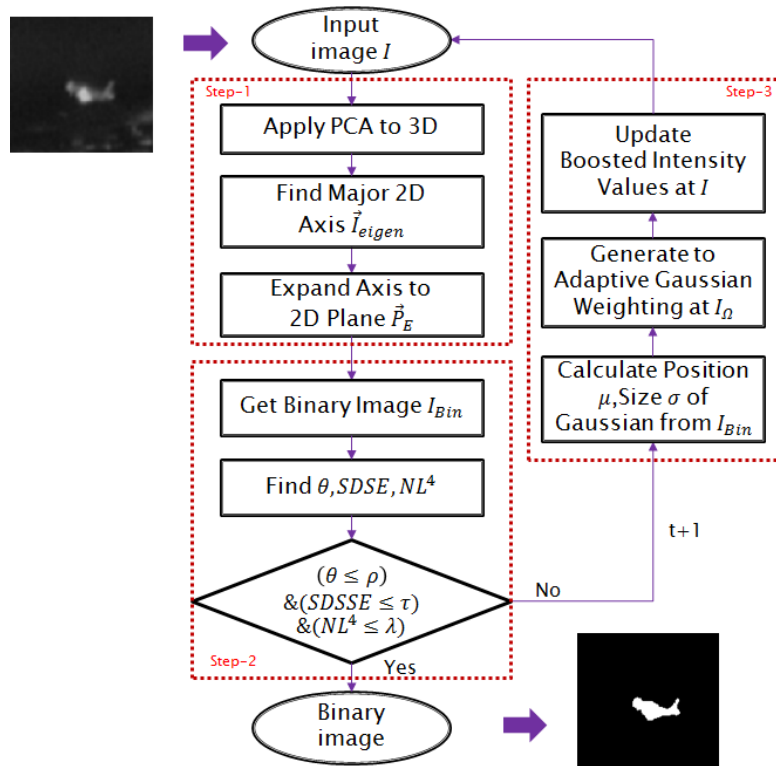


Fig. 1. Flow chart of the proposed method.

Therefore, an analysis technique to model both the target and background regions is required. To this end, the PCA method can be utilized, where the minimized TLS error data is used as a modeling technique by considering the observational noise errors of both the dependent and independent variables. In a PCA, blind noises are considered to be in both the target and background [13]. Moreover, PCA uses a dimension reduction technique based on the minimized errors method for separating the target region from the non-homogeneous backgrounds. The model of the target with a non-homogeneous background region can be simply expressed as Eq. (1).

$$I[x, y, z]^T = I_T + I_B + \varepsilon \quad (1)$$

2.1 PCA technique in the IR image by using minimized TLS errors

In order to separate the target region I_T and the non-homogeneous backgrounds I_B , we apply the PCA method based on the minimized TLS errors \mathcal{E} to the 3D intensity space from IR as Eq. (2).

$$\vec{P}_E[x, y, z]^T = (\alpha \times \vec{I}_{eigen}) + \vec{I}_{mean}, \quad \vec{I}_{eigen} = PCA_{eig}(\vec{I}), \quad \alpha = \max(\vec{I}_w, \vec{I}_h) \quad (2)$$

Consequently, in Eq. (2), an Eigen vector of 2D plane \vec{I}_{eigen} is generated to distinguish the target region from a non-homogeneous environment by using PCA method with minimized TLS errors values from the 3D intensity space. **Fig. 2(a)** and **(b)** show the initial result of target segmentation by using a PCA plane. The data of below the PCA plane are removed in the 3D space from previous step. As shown in **Fig. 2(b)**, \vec{P}_E can be extended as the plane function multiplied by the scalar value α . **Fig. 2(c)** shows that only the data of above than PCA plane is used to calculate the weight data for next step. In Eq. (2), α is the largest value among the maximum value of the width \vec{I}_w and height \vec{I}_h from the input image. Moreover, a \vec{I}_{mean} value is added for determining the height of the plane in 3D intensity space.

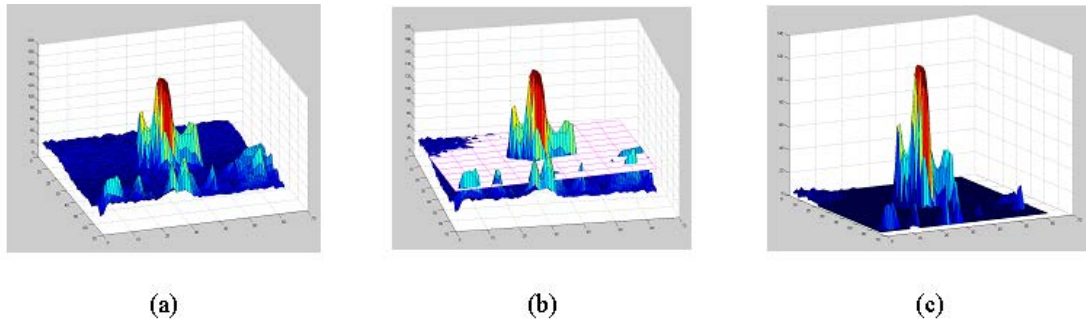


Fig. 2. The initial result of target segmentation by using a PCA plane: (a) 3D Plot, (b) PCA plane in 3D, (c) Data above PCA plane.

2.2 Save the pixels with higher values than the plane

In order to estimate the initial non-homogeneous environment, we determine pixels \hat{I}_Ω of the generated plane above as target region by binary result I_{Bin} by using Eq. (3). As shown in **Fig. 3**, a binary image I_{Bin} is generated by using Eq. (3).

$$I_{Bin} = \begin{cases} 1 & \text{if } \hat{I}_\Omega \\ 0 & \text{otherwise} \end{cases}, \quad \hat{I}_\Omega = \vec{I} \geq \vec{P} \quad (3)$$

2.3 Angle between the brightness plane and the PCA vector with SDSSE

In order to estimate the degree of non-homogeneous environment, we can obtain the angle θ between the PCA plane and the brightness plane \vec{B} in the intensity domain by using Eq. (4).

$$\theta = \tan \left(\frac{\vec{I}_{eigen} \times \vec{B}}{\sqrt{(x^2 + y^2 + z^2)}} \right)^{-1} \times \frac{180}{\pi}, \quad \vec{B} = [x_b = 0, y_b = 0, z_b = 1]^T \quad (4)$$

To more effectively segment the target region, the sum of the square error (SSE) is calculated in Eq. (5).

$$SSE = \sum_{i=1}^{W \times H} (\bar{P}_i - \bar{I}_i)^2 \quad (5)$$

Furthermore, the second derivative of the sum of the square error (SDSSE) is calculated by using the absolute of SSE value from the previous step of $t-1$ in Eq. (6), and t is the meaning of the current step numbers.

$$SDSSE_t = \frac{d^2(SSE_{t-1})}{dt^2} \quad (6)$$

2.4 Apply in adaptive Gaussian weighting kernel and save the labeling numbers

In order to boost the intensity level of the target region, we generate a Gaussian weighting function $f_{gauss}(x, y)$ at center of mass position x, y in binary image. However, it is difficult to determine the size and position of the adaptive Gaussian kernel window exactly. Since the intensity values of the target region are mixed with the background intensity values. To overcome this problem, we are determining the size of the Gaussian kernel windows from the binary image for boost the intensity level of target region.

As shown in [Fig. 3](#), the position and size of the Gaussian kernel was determined by using the mean $\mu_{x,y}$ and standard variation $\sigma_{x,y}$ values from the binary image by using Eq. (7).

Furthermore, the labeling number NL^4 is obtained by using 4-neighbor label connected components in the binary result of the target region.

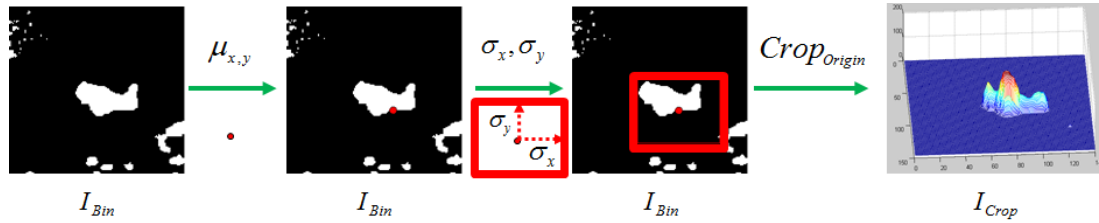


Fig. 3. Cropped image by using the initial binary image.

And then, the region of the target for the cropping image is generated by using the mean and standard variation $\sigma_{x,y}$ values. As a result, the cropped image is added by the Gaussian weight function for boosting the intensity value as in Eq. (7).

$$f_{gauss}(x, y) = \exp\left(-\left(\frac{(x - \mu_x)^2}{2\sigma_x^2} + \frac{(y - \mu_y)^2}{2\sigma_y^2}\right)\right) \tag{7}$$

However, generally, the total sum of the Gaussian weight value is 1 as shown in Eq. (7). Therefore, we multiply at Gaussian weight function by mean value $mean(I_{Crop})$ of cropped image for using weighting kernel K by Eq. (8). The Gaussian type weight is then added to the cropped region from the input image in Eq. (8).

$$K = I_{Crop} + (f_{gauss} \times mean(I_{Crop})) \tag{8}$$

To apply to the ROI region image, the cropped image is required for normalization. To this end, the $K_{x,y}^{Normalize}$ is guaranteed between the minimum brightness value K_{min} and the maximum value K_{max} by using Eq. (9).

$$K_{x,y}^{Normalize} = \frac{K_{x,y} - K_{min}}{K_{max} - K_{min}} \times K_{max} \tag{9}$$

Finally, the previously calculated Gaussian kernel is updated in the input image by weighted $K_{x,y}^{Normalize}$ in the I_{Ω} region. In **Fig. 4** shows that the Gaussian weighted kernel is updated at cropped region from the previous step. **Fig. 5** shows that the cropped region with the input image is the updated value by the adaptively weighted Gaussian type kernel to clarify the convergence criteria. Therefore, the propose method can get a more effective PCA plane by the adaptively updated Gaussian weighting kernel.

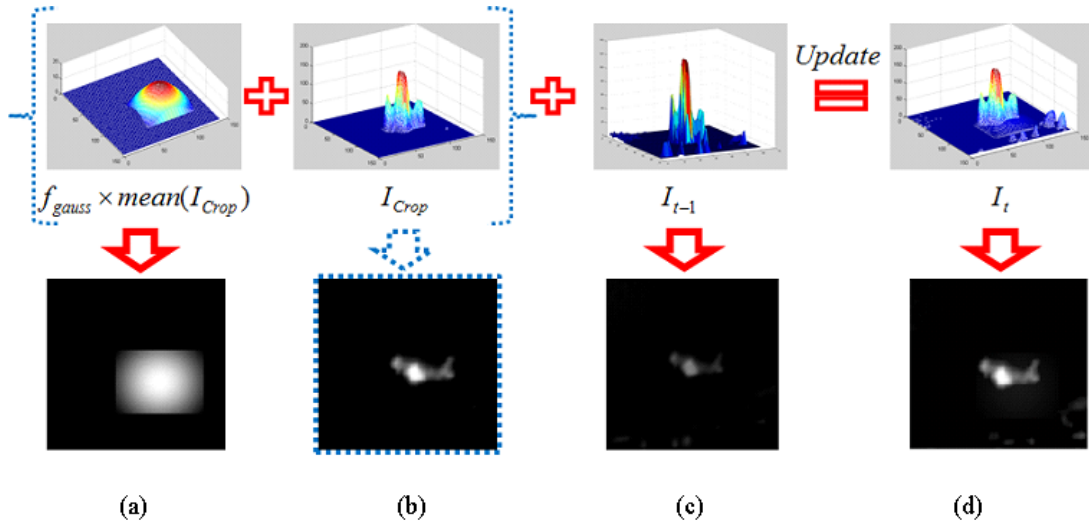


Fig. 4. The result of Gaussian weight: (a) Kernel of Gaussian, (b) Normalized image, (c) Before applying Gaussian weight, (d) After applying Gaussian weight.

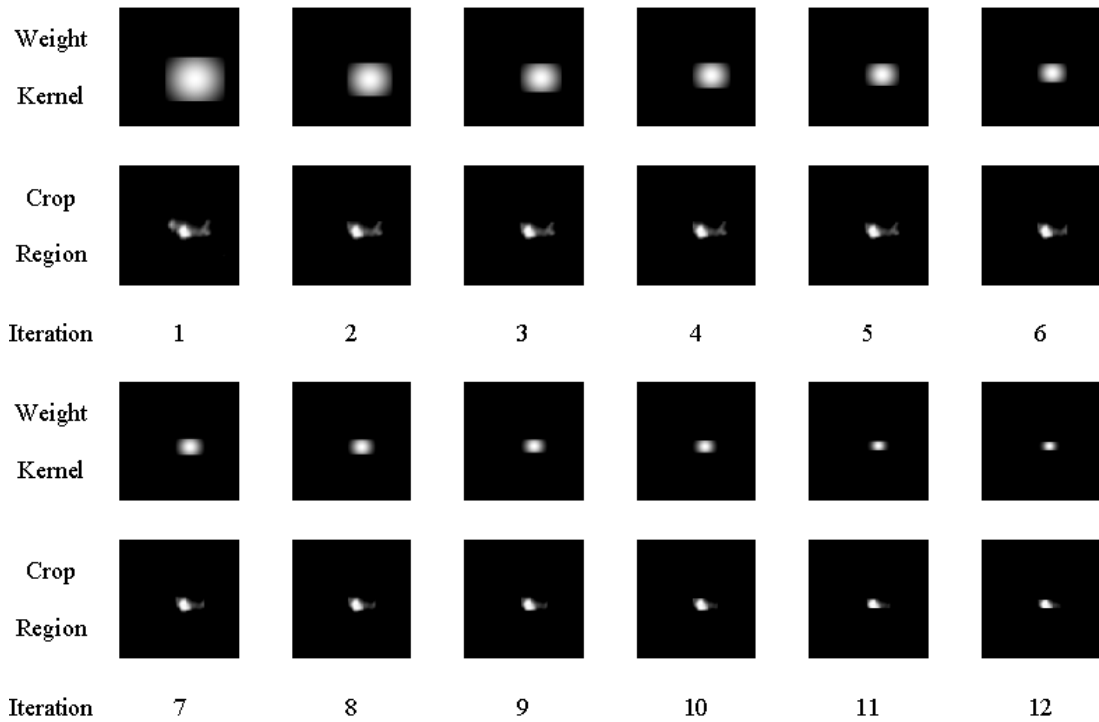


Fig. 5. Weighted Gaussian kernel and cropped region image.

2.5 Segmentation based on the plane by repeated subtraction of the previous step until convergence

In order to effectively separate the target region, threshold values are used for convergence conditions. The threshold values consist of the SDSSE, the angle and labeling values. As a result, the region of the target is more clearly separate from the non-homogeneous

environments by using the threshold values. The convergence condition value TH_t is iteratively updated until convergence in Eq. (10), and ρ, τ, λ is the error constant value.

$$TH_t = (\theta_t \leq \rho) \& (SDSSE_t \leq \tau) \& (NL_t^4 \leq \lambda) \quad (10)$$

Therefore, when do not meet the convergence condition value TH_t , the adaptive Gaussian kernel is the weighted addition in the PCA plane with the non-removed data from the previous step. Consequently, the data that do not meet the criteria are removed.

As shown in From **Fig. 2(a)** to **Fig. 2(c)**, the data \vec{I}_{Blow} of below the PCA plane \vec{P}_t are removed in the 3D space \vec{I}_t from previous step \vec{I}_t . Then only the data \vec{I}_{t+1} of above than PCA plane \vec{P}_t is used to calculate the weight data for next step as Eq. (11). Consequently, the value of the above data in PCA plane was used for the next step.

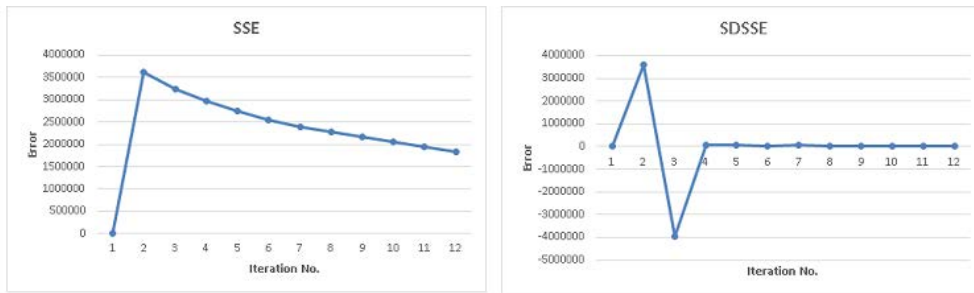
$$\vec{I}_{t+1} = \vec{I}_t - (\vec{I}_{Blow}), \quad \vec{I}_{Blow} = \vec{I}_t < \vec{P}_t \quad (11)$$

Table 1 and **Fig. 6** show the example of change of the SSE, SDSSE and θ values in each step by using the proposed method. **Fig. 7** shows an example result image of each step by using proposed method. As you can see, the result of segmentation image is repetitions until the criteria is met by the adaptively Gaussian weighted kernel values at the target region at each step.

Table 1. Values of SSE, SDSSE and Theta at each step.

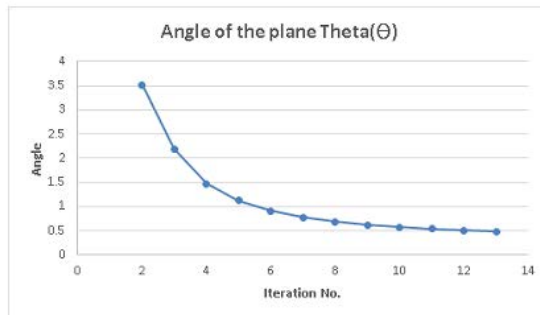
Step	SSE	SDSSE	Theta(θ)
1	5.4430	0	3.5219
2	3.61E+06	3.61E+06	2.1934
3	3.25E+06	3.97E+06	1.4758
4	2.97E+06	8.00E+04	1.1216
5	2.75E+06	6.00E+04	0.9069
6	2.55E+06	2.00E+04	0.7806
7	2.40E+06	5.00E+04	0.6859

8	2.28E+06	3.00E+04	0.6217
9	2.16E+06	0.00E+00	0.5794
10	2.05E+06	1.00E+04	0.5436
11	1.94E+06	0.00E+00	0.5117
12	1.84E+06	1.00E+04	0.4824



(a)

(b)



(c)

Fig. 6. SSE, SDSSE and Theta at each iteration: (a) Sum of the Square Error, (b) Second derivative of sum of the square error, (c) Angle of the plane Theta.

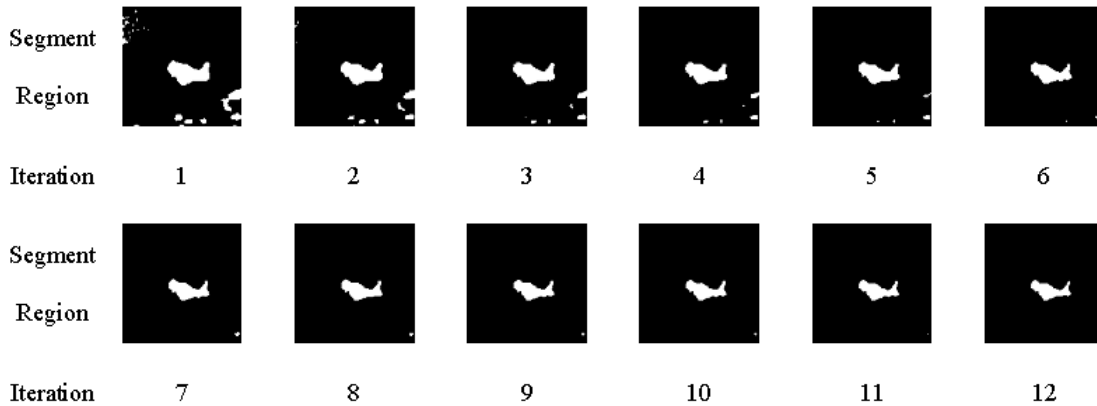


Fig. 7. The result of segmentation at each step.

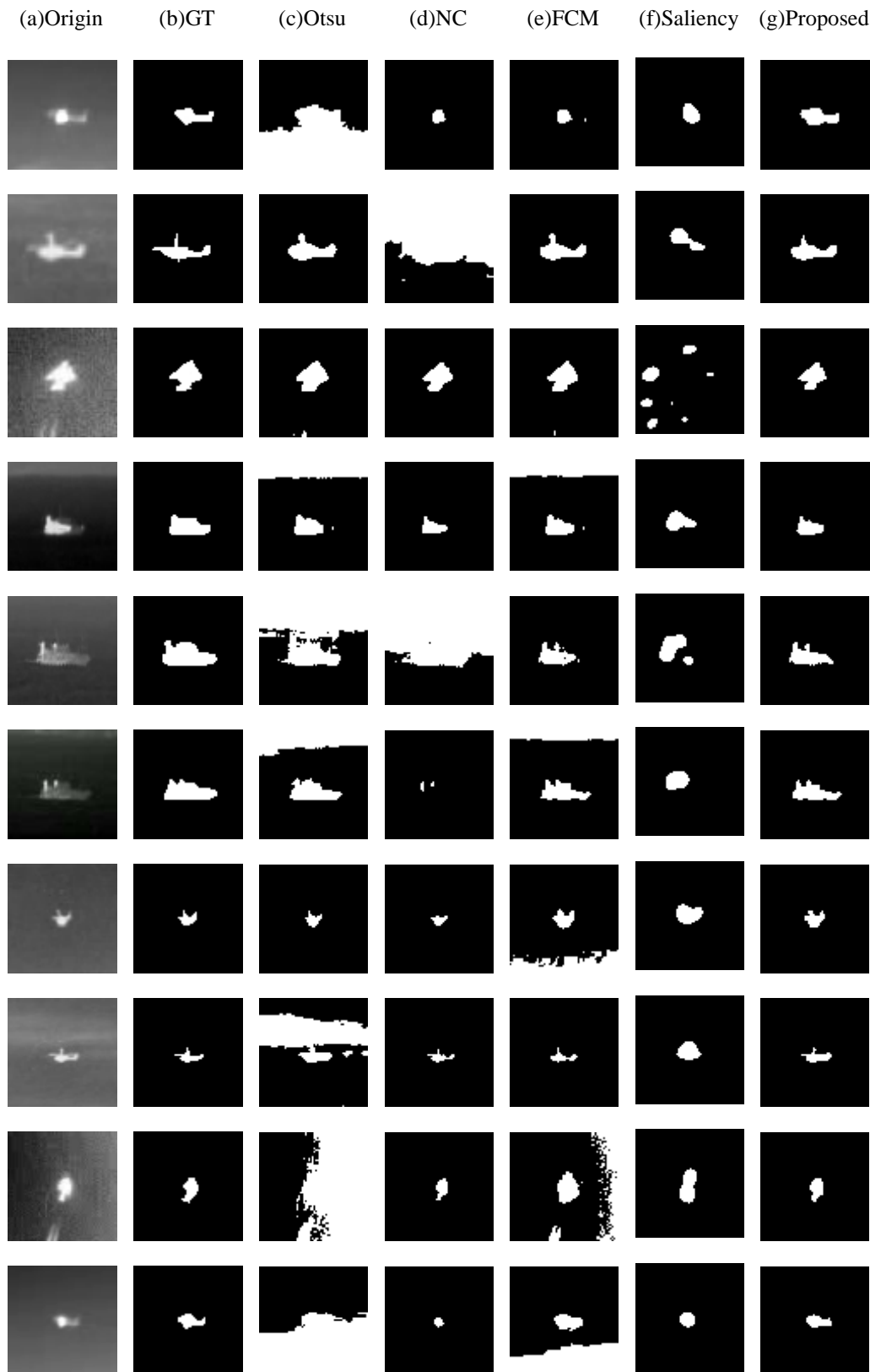
3. Experimental results

The proposed method was tested and its performance compared to the conventional methods by using a test set composed of 100 IR images containing aircraft and ship targets under cluttered background environments. The test images with pixel resolution of 640 x 480 were obtained from a mid-wave infrared camera of FLIR systems. Input ROI images were presented with a size of 64 x 64. In this paper, we set the parameter value as $\rho = \tau = 0.00001$, and $\lambda = 2$.

The error constant value when the ρ and τ has low value, non-homogeneous characteristic has low value. Therefore, in this paper, we iteratively apply to segmentation target region by using characteristic of non-homogeneous with error constant value of ρ , τ . As shown in **Fig 6.(b)** and **Fig 6.(c)**, the error constant value ρ , τ is close to zero value. However, these values never have zero value during iteration steps. In order to solve these problems, we adopt an experimentally reasonable error constant value until the convergence criteria are met.

The reasonable error constant value was experimentally set to 0.00001 for close to 0. And the constant value λ is the number of clustered labeling result by using 4-neighbor connected component in the binary result image. The target region should have one set labeling result. However, the experimental results sometimes show that the target region has a more separated segmentation region than one region. In order to stop the iteration process, the reasonable error constant value was experimentally set to 2 for close to 1.

Fig. 8 shows the experimental results of the proposed method and the conventional methods. **Fig. 8(a)** shows the input ROI images obtained from the infrared images. **Fig. 8(b)** shows the manually extracted ground truth. From **Fig. 8(c)** to **Fig. 8(f)**, the results are obtained by using conventional methods including Otsu's method, normalized cuts, fuzzy c-means and saliency method. **Fig. 8(g)** shows the results of the proposed method. As shown in **Fig. 8**, the results of the proposed method are more robust against irregular intensity environments compared to the conventional methods. In order to quantitatively evaluate the target extraction performance, we utilize the pixel-based quality measure [14] to calculate the extraction error rate, as shown in Eq. (12).



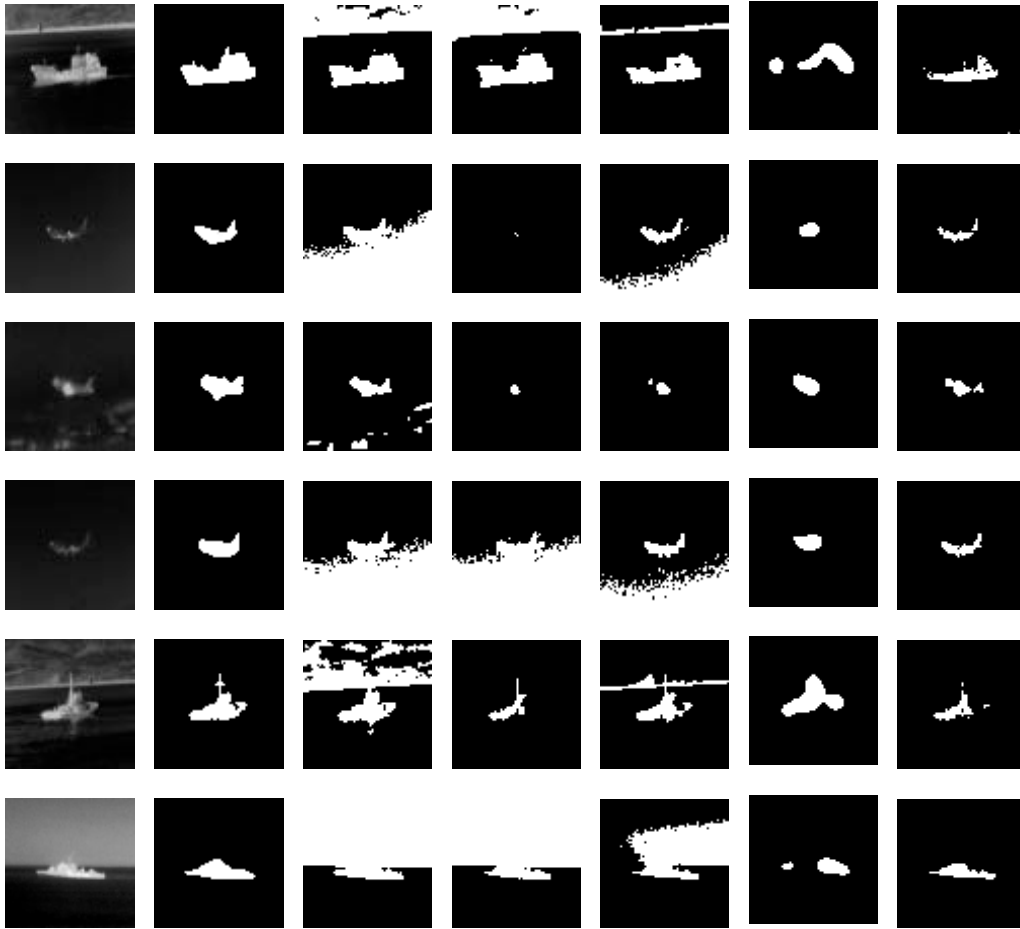


Fig. 8. Comparisons of the proposed method and the conventional methods: (a) Input image, (b) Ground truth, (c) Otsu's method, (d) Normalized cut method (e) FCM method, (f) Saliency method, (g) The proposed method.

$$E(B_{img}, B_{gt}) = \frac{\sum_{(x,y)} B_{img}(x,y) \otimes B_{gt}(x,y)}{\sum_{(x,y)} B_{gt}(x,y)} \quad (12)$$

where B_{img} and B_{gt} are the binary image for the extracted target and the manually extracted ground truth, respectively, and \otimes denotes the exclusive OR operation. **Table 2** shows 100 images of the performance comparisons in terms of the extraction error rate calculated from the results of Otsu's method, normalized cuts, fuzzy c-means, saliency method and the proposed method. By evaluating the experimental results, the proposed method achieves much better performance than conventional methods. In **Table 2**, the meaning of Avg. and Std. Dev. are the average value and the standard deviation of the extraction error rate, respectively. The meaning of min and max are the minimum and the maximum value of the extraction error rate, respectively. In terms of the average error rate, the proposed method achieves much better performance than conventional methods, as shown in **Table 2**. The average computational time tested on 100 images for Otsu, Normalized cuts, Fuzzy c-means, Saliency method and

our method are shown in **Table 3**. It is obvious from **Table 3** that our method is very time efficient compared to the Normalized cuts and Fuzzy c-means methods. The algorithm is tested on a Quad Core desktop of 3.07GHz, 4GB DDR.

Table 2. Performance comparison in terms of extraction error rate.

Method	Extraction error rate			
	Avg.	Std. Dev.	Min	Max
Proposed method	0.38	0.24	0.08	1.67
Otsu	6.07	8.58	0.17	35.22
Normalized cuts	2.20	3.76	0.13	19.93
Fuzzy c-means	2.35	4.37	0.17	31.98
Saliency method	1.04	1.03	0.38	7.43

Table 3. Time complexity comparison of each algorithm.

Method	Otsu	Normalized cuts	Fuzzy c-means	Saliency method	Proposed
Seconds	0.0060	0.2779	0.4197	0.0299	0.0411

4. Conclusion

In this paper, we propose a novel method to segment the exact target region in an IR image. This method is necessary because conventional approaches for extracting targets have only utilized the intensity values of pixels with the analysis targets of the intensity values. However, the intensity values of the target region are mixed with the background region in a non-homogeneous environment. Therefore, it is difficult to extract the target regions effectively due to the non-homogeneous environment. To this end, minimized TLS errors based on PCA and adaptive Gaussian weight kernel were utilized for segmenting the target region. The experimental results show that the proposed method achieves better performance of the target segmentation than the Otsu's, Normalized cut, FCM and Saliency method.

References

- [1] E. Vasquez, F. Galland, G. Delyon, and P. Refregier, "Mixed segmentation-detection-based technique for point target detection in nonhomogeneous sky," *Appl. Opt.* Vol. 49, pp. 1518-1527, 2010. [Article \(CrossRef Link\)](#)
- [2] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. on Systems,*

- Man, and Cybernetics*, Vol. 9, No. 1, pp. 62-66, Jan., 1979. [Article \(CrossRef Link\)](#)
- [3] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence* 22(8), 888-905, 2000. [Article \(CrossRef Link\)](#)
- [4] W. Tao, H. Jin, and L. Liu, "A new image thresholding method based on graph cuts," in *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 605-608, 2007. [Article \(CrossRef Link\)](#)
- [5] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms," *Plenum Press, New York*, 1981. [Article \(CrossRef Link\)](#)
- [6] M.M. Trivedi and J. C. Bezdek, "Low-level segmentation of aerial images with fuzzy clustering," *IEEE Trans. Syst., Man Cybernet.* 16(4), 589-598, 1986. [Article \(CrossRef Link\)](#)
- [7] J. Wu, J. Li, J. Liu, and J. Tian, "Infrared image segmentation via fast fuzzy c-means with spatial information," in *Proc. of IEEE Int. Conf. on Robotics and Biomimetics*, pp. 742-745, 2004. [Article \(CrossRef Link\)](#)
- [8] X. Hou, J. Harel, and C. Koch, "Image Signature: Highlighting Sparse Salient Regions," *IEEE Trans. Pattern Anal. Mach. Intell.* 34(1): 194-201, 2012. [Article \(CrossRef Link\)](#)
- [9] R. C. Gonzalez and P. Wintz, "Digital Image Processing," Addison-Wesley Publishing Company, 1987.
- [10] H. D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *Image Processing, IEEE Transactions on* Vol. 9, Issue 12, pp. 2071-2082, Dec. 2000. [Article \(CrossRef Link\)](#)
- [11] F. D. Fengzhi, Z. Jinbin, Z. Huailin, "The application of resonance algorithm for image segmentation," *Applied Mathematics and Computation*, Vol. 194, Issue 2, 15 Dec. 2007, pp. 453-459, ISSN 0096-3003, 10.1016/j.amc.2007.04.047. [Article \(CrossRef Link\)](#)
- [12] S. Ji, Y. Wu, Y. Wu, "Infrared small target detection based on complex contourlet transform and principal component analysis," in *Proc. of Image and Signal Processing (CISP), 2010 3rd International Congress on*, Vol.6, No., pp.2579-2583, 16-18 Oct. 2010. [Article \(CrossRef Link\)](#)
- [13] T. G. Davis, "Total least squares spiral curve fitting," *Journal of Surveying Engineering*, Vol. 125, No. 4, pp. 159-176, 1999. [Article \(CrossRef Link\)](#)
- [14] Z. Liu, W. Li, L. Shen, Z. Han, and Z. Zhang, "Automatic segmentation of focused objects from images with low depth of field," *Pattern Recognition Letters* 31(7), 572-581, 2010. [Article \(CrossRef Link\)](#)



Yong Min Kim received his MS degree in computer science and engineering from Hanyang University, Republic of Korea, in 2011. He is currently working toward his PhD degree at the computer vision and pattern recognition laboratory, department of computer science and engineering, Hanyang University, Republic of Korea. His major interests include face processing, object segmentation, and target tracking.

Ki Tae Park received the BS, MS, and PhD degrees in computer science from Hanyang University, Republic of Korea, in 2000, 2002, and 2007, respectively. He had worked in Samsung Electronics Co., Ltd., Suwon, Republic of Korea. And He was worked as research of professor, Hanyang University, Republic of Korea, from 2009 to 2014. He is currently working at the attached institute of ETRI. His major research interests are in the areas of digital signal processing, content-based image retrieval, object segmentation, multimedia processing, and video indexing.



Young Shik Moon is the corresponding author of this paper. He received the B.S. and M.S. degrees in electronics engineering from Seoul National University and Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1980 and 1982, respectively, and Ph.D. degree in electrical and computer engineering from the University of California at Irvine, CA, in 1990. From 1982 to 1985, he had been a researcher at the Electronics and Telecommunications Research Institute, Daejeon, Korea. In 1992 he joined the department of Computer Science and Engineering at Hanyang University, Korea, as an Assistant Professor, and is currently a Professor. His research interests include computer vision, image processing, pattern recognition, and computational photography. Dr. Moon served as General Chair of 2014 IEEE International Symposium on Consumer Electronics, and he is now the President of the Institute of Electronics and Information Engineers, Korea