Multi-Level Segmentation of Infrared Images with Region of Interest Extraction

Seokwon Yeom
School of Computer and Communication Engineering, Daegu University, Gyeongsan, Korea

Abstract

Infrared (IR) imaging has been researched for various applications such as surveillance. IR radiation has the capability to detect thermal characteristics of objects under low-light conditions. However, automatic segmentation for finding the object of interest would be challenging since the IR detector often provides the low spatial and contrast resolution image without color and texture information. Another hindrance is that the image can be degraded by noise and clutters. This paper proposes multi-level segmentation for extracting regions of interest (ROIs) and objects of interest (OOIs) in the IR scene. Each level of the multi-level segmentation is composed of a $k$-means clustering algorithm, an expectation-maximization (EM) algorithm, and a decision process. The $k$-means clustering initializes the parameters of the Gaussian mixture model (GMM), and the EM algorithm estimates those parameters iteratively. During the multi-level segmentation, the area extracted at one level becomes the input to the next level segmentation. Thus, the segmentation is consecutively performed narrowing the area to be processed. The foreground objects are individually extracted from the final ROI windows. In the experiments, the effectiveness of the proposed method is demonstrated using several IR images, in which human subjects are captured at a long distance. The average probability of error is shown to be lower than that obtained from other conventional methods such as Gonzalez, Otsu, $k$-means, and EM methods.

Keywords: Infrared image segmentation, Region of interest extraction, Multilevel segmentation, Gaussian mixture modeling, Statistical image processing

1. Introduction

Infrared (IR) image processing has been utilized for surveillance, robotic systems, intelligent vehicles, biometric recognition, and non-destructive monitoring. IR radiation is capable of capturing scenes that are not visible at night and in bad weather. However, IR images provide no color and texture information, and the image quality is often degraded by low spatial and contrast resolution as well as noise and clutters.

In the past, various researches have been performed to take advantage of thermal imaging under non-visible conditions. Region of interest (ROI) extraction with a one-bit transform has been proposed for a single image in [1]. The IR target extraction has been researched based on the Gaussian mixture distributions with horizontal and vertical projection modeling [2]. In [3], a two-dimensional histogram method has been proposed considering the location and intensity of the object. The segmentation by the expectation-maximization (EM) clustering for IR images has been evaluated in [4]. Low-resolution person detection by hot spot classification
was developed for moving targets, which requires the image sequence rather than a single frame [5]. The ROI based target detection with texture features has been proposed in [6]. A target extraction method with fuzzy thresholding has been studied in [7]. A novel approach was presented in [8] that can help to determine the pedestrian’s orientation in the IR scene. A two-step detection and tracking method, which is based on the support vector machine (SVM) and Kalman filtering was proposed in [9]. Real-time pedestrian detection and tracking at nighttime has been demonstrated through two-stage detection and template-matching-based tracking [10]. In [11], a shape-independent method with far-infrared imaging has been addressed. A correlation-based regional template matching algorithm has been developed for multiple target detection and tracking [12]. In [13], the human subjects are segmented without considering the ROI extraction. Edge-based segmentation using histogram of oriented gradient (HOG) with the ship size ratio is proposed in [14].

In this paper, the multi-level segmentation is addressed for extracting ROIs as shown in Figure 1(a). The foreground objects are segmented from the individual ROI windows by the EM algorithm. Figure 1(b) shows the multi-level segmentation with L levels. Each level comprises a k-means clustering algorithm, an EM algorithm, and a Bayesian decision rule. The histogram of the IR scene is assumed to follow the Gaussian mixture distribution. The k-means clustering algorithm initializes the Gaussian mixture model (GMM) parameters, namely, the means, variances, and weights. The EM algorithm iteratively estimates the parameters until convergence [15]. The Bayesian decision rule assigns each pixel to a cluster maximizing the posterior probability. Only extracted cluster regions at one level are considered the input to the next level segmentation, thus the segmentation process is consecutively performed narrowing the area in the image.

The multi-level method has proved their effectiveness by analyzing passive millimeter wave images [16], where only two segmentation levels are considered with the prior knowledge of the ROI. However, in this paper, the multi-level segmentation method has been extended to segment the multiple ROI windows which can be located at any arbitrary position.

In the experiments, several IR images are processed where a number of human subjects are captured at a long distance. They are shown to be corrupted by noise and clutter. The proposed method shows a lower average error probability than that of other methods such as Gonzalez, Otsu, k-means, and EM methods.

The paper is organized as follows: In Section 2, the multi-level segmentation and the performance evaluation metric are described. In Section 3, the experimental results are demonstrated. The conclusion follows in Section 4.

2. Multilevel Segmentation for ROI Extraction

The multi-level process comprises of several segmentation levels. Each level is composed of the k-means clustering algorithm with splitting initialization, the EM algorithm with the GMM model, and the Bayesian decision rule.

2.1 k-means Algorithm with Splitting Initialization

The k-means clustering algorithm initializes the parameters of the Gaussian mixture distribution [17]. Figure 2 shows the k-means procedure to generate two sets (clusters); $S_+$ and $S_-$ from a set $S = \{x_1, ..., x_N\}$; $x_i$ corresponds to the intensity at the i-th pixel in the image. The initial value $y_0$ is the mean
value of the set $S$, $\Delta$ is an arbitrary value for initial splitting, and $n_p$ is the number of pixels in the image. Two clusters, $S_+$ and $S_-$ correspond to the segmented areas. This process can be repeated when more than two clusters are required. After the $k$-means algorithm, the GMM parameters are initialized by sample means, sample variances, and the proportions of the clusters.

### 2.2 EM Algorithm and Decision Process

The histogram of the IR scene is modeled with a Gaussian mixture distribution as follows:

$$p(x_j) = \sum_{k=1}^{n_k} P(G_k | x_j) = \sum_{k=1}^{n_k} N(x_j | \mu_k, \sigma_k) P(G_k), \ j = 1, ..., n_p,$$

where $N$ represents the Gaussian distribution, $G_k$ denotes the hypothesis of the cluster $k$, $\mu_k$ and $\sigma_k$ are the mean and the variance of the cluster $k$, respectively, and $n_k$ is the number of clusters. The EM algorithm iteratively estimates the parameters, $\mu_k$, $\sigma_k$ and $P(G_k)$, $k = 1, ..., n_k$ as illustrated in Figure 3; $i$ represents the number of iteration, $\varepsilon$ is the termination criterion, and $L_i$ is the sum of log-likelihood: $L_i = \sum_{j=1}^{n_p} \log p_i(x_j)$.

More detailed procedures of the EM algorithm can be found in [15]. Each pixel is assigned to one of clusters by the following decision rule:

$$\hat{k}_j = \arg \max_{k=1, ..., n_k} P(G_k | x_j), \ j = 1, ..., n_p. \quad (2)$$

### 2.3 Performance Evaluation

The average probability of error is chosen for performance evaluation [16, 18] as

$$P_e = P(D_1 | H_0) P(H_0) + P(D_0 | H_1) P(H_1) = \frac{|F_S - F_O| + |F_O - F_S|}{|I|}, \quad (3)$$

where $H_1$ and $H_0$ indicate the foreground object and the background hypothesis, respectively, and $D_1$ and $D_0$ indicate the decision for the object and the background region, respectively; $F_S$, $F_O$, and $I$ indicate the pixel sets of the segmented object, the ground truth region of the object, and the entire image, respectively, and $\mid \cdot \mid$ denotes the number of pixels in the set. It is manifest that $I = B_O \cup F_O$, where $B_O$ is the ground truth region of the background. Figure 4 demonstrates the segmented area and the ground truth for the calculation of the average probability error.
3. Experimental Results

In the experiments, four IR images in Figure 5 are processed by the multi-level segmentation method. Several human subjects are captured in the IR images at a long distance. The image sizes of Figures 5(a)-5(d) are $240 \times 360$, $240 \times 360$, $240 \times 320$, and $192 \times 256$ pixels, respectively.

Figure 6(a) is the histogram of Figure 5(a) and the GMM fitting after the first level segmentation. The white areas in Figure 6(b) correspond to the blue line Gaussian distribution in Figure 6(a). The white areas in Figure 6(b) become the input to the second level segmentation resulting in Figure 6(c) and 6(d). The white areas in Figure 6(d) correspond to the right side (red line) Gaussian distribution in Figure 6(c). The extracted white areas in Figure 6(d) are input to the third level segmentation. Figure 6(e) and 6(f) are the third level segmentation results. The white areas in Figure 6(f) correspond to the right side (red line) Gaussian distribution in Figure 6(e).

Figure 7(a) shows four ROI windows extracted from Figure 6(f). The blue circles are the centroids of the windows. The ROI window is set 1.5 times bigger than the minimum rectangle including the segmented areas. The windows which are less than 50 pixels are discarded assuming the minimum size of the object is known. Each ROI window is processed individually by the EM method to segment the foreground object. Figure 7(b)-7(d) show the first ROI window, the histogram of Figure 7(b), and the human subject segmented from Figure 7(b), respectively. The white area in Figure 7(d) corresponds to the right side (red line) Gaussian distribution in Figure 7(c). Figure 7(e) is the combined segmentation result from all four ROI windows. Figure 8(a)-8(c) show the ROI extraction, ROI windows, and final segmentation result of Figure 5(b)-5(d), respectively.

Figure 9 shows the segmentation results of IR scenes 1-4 with other methods: EM, $k$-means, Otsu and Gonzalez from left to right column. Table 1 shows the average probability error in Eq. (3). The ground truth region is manually decided from the original IR scenes. It is shown that the average probability error
of the multi-level method is more than 50% lower on average than that of other methods.

4. Conclusion

In this paper, the multi-level segmentation is addressed for the ROI and the object of interest (OOI) extraction of the IR scene. The histogram of the image is assumed to follow the Gaussian mixture distribution. The multi-level segmentation has the several stages composed of the $k$-means clustering algorithm, the EM method, and the decision rule. It provides an effective way to extract the ROIs in the noisy scene. It has been shown that the segmentation of the human subjects becomes more accurate when the ROIs obtained are individually processed.
Further investigation on the classification of the segmented object remains for future study.

**Conflict of Interest**

No potential conflict of interest relevant to this article was reported.

---

**Acknowledgements**

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2012R1A1A2008545).

**References**

[1] S. Erturk, “Region of interest extraction in infrared images using one-bit transform,” *IEEE Signal Processing Letters*, vol. 20, no. 10, pp. 952-955, 2013. [http://dx.doi.org/10.1109/LSP.2013.2274637](http://dx.doi.org/10.1109/LSP.2013.2274637)

[2] D. K. Shin and Y. S. Moon, “Extraction of infrared target based on Gaussian mixture model,” *IEEE Transactions*
Figure 9. Segmentation by EM, k-means, Otsu, Gonzalez method from left to right column, (a) IR Sene 1, (b) IR Sene 2, (c) IR Sene 3, (d) IR Sene 4.
[15] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press, 1995.

[16] S. Yeom, D. S. Lee, J. Y. Son, M. K. Jung, Y. Jang, S. W. Jung, and S. J. Lee, “Real-time outdoor concealed-object detection with passive millimeter wave imaging,” *Optics Express*, vol. 19, no. 3, pp. 2530-2536, 2011. http://dx.doi.org/10.1364/OE.19.002530

[17] A. Gersho and R. M. Gray, “Predictive quantization,” in *Vector Quantization and Signal Compression*, A. Gersho and R. M. Gray, Eds. Boston, MA: Springer, 1992, pp. 203-223. http://dx.doi.org/10.1007/978-1-4615-3626-0_7

[18] W. A. Yasnoff, J. K. Mui, and J. W. Bacus, “Error measures for scene segmentation,” *Pattern Recognition*, vol. 9, no. 4, pp. 217-231, 1977. http://dx.doi.org/10.1016/0031-3203(77)90006-1

[19] J. W. Davis and M. A. Keck, “A two-stage template approach to person detection in thermal imagery,” in *Proceedings of Seventh IEEE Workshops on Application of Computer Vision*, Breckenridge, CO, 2005. http://dx.doi.org/10.1109/ACVMOT.2005.14

Seokwon Yeom is currently an associate professor in the School of Computer and Communication Engineering at Daegu University in South Korea. He was a visiting scholar at University of Maryland in 2014. He received the Ph.D. degree in Electrical and Computer Engineering from the University of Connecticut in 2006 and the M.S. and B.S. degrees in Electronics Engineering from Korea University, South Korea. His research interests are intelligent signal and image processing, and optical information processing. He is now performing several research projects granted by Korea government and industries. Research area: intelligent signal and image processing. optical information processing. E-mail: yeom@daegu.ac.kr

[20] IEEE OTCBVS WS Series Bench and R. Miezianko, “Terravic research infrared database,” Available http://vcipl-okstate.org/pbvs/bench/