Pedestrian identification in infrared images using visual saliency detection technique

Mai Thanh Nhat Truong*, Sanghoon Kim*
*Department of Electrical, Electronic, and Control Engineering, Hankyong National University

Abstract

Visual saliency detection is an important part in various vision-based applications. There are a myriad of techniques for saliency detection in color images. However, the number of methods for saliency detection in infrared images is inadequate. In this paper, we introduce a simple approach for pedestrian identification in infrared images using saliency. The input image is thresholded into several Boolean maps, an initial saliency map is then calculated as a weighted sum of created Boolean maps. The initial map is further refined by using thresholding, morphology operation, and Gaussian filter to produce the final, high-quality saliency map. The experiment showed that the proposed method produced high performance results when applied to real-life data.

1. Introduction

Saliency detection research plays an important role in computer vision, because visual saliency is crucial information in various vision-based applications such as content-aware image manipulation [1], video compression [2], image segmentation [3], and object recognition [4]. Visual saliency detection is a classification task on pixel-level, in which each pixel is assigned a saliency score to determine whether it belongs to the foreground (salient objects) or the background. Saliency detection techniques can be divided into two major categories: bottom-up and top-down. Bottom-up approaches utilize low-level features such as color, texture, orientation, while top-down approaches depend on high-level factors such as prior knowledge regarding tasks or events.

Borji et al. [5] performed evaluations on 41 state-of-the-art models with seven data sets. The report provided an in-depth analysis for the development of saliency research over the last few years. Another popular benchmark for evaluating saliency detection methods is MIT Saliency Benchmark [6] which contains results from over 70 models. Unfortunately, techniques for saliency detection in infrared images are not as common as in color images. Infrared imaging is an specialized technology which has various applications in critical fields such as security, surveillance, and military [7]. Saliency detection algorithms that work well with color images may show poor performance when applied to infrared images. It is because infrared images are usually highly noisy, low contrasts, besides they have low resolutions and insufficient features.

Several attempts have been made to deal with these difficulties. Li et al. [8] used Gaussian scale-space representation and signal theory to compute phase information for the purpose of detecting both bright and dark salient objects. In [9], Li et al. first improved the quality and contrast of the image in frequency domain, then they combined luminance distribution and gradient feature to construct the saliency map. Qin et al. [10] took advantage of human vision-based attention mechanism and information theory to create a full resolution saliency image. In this research, we utilize Boolean maps [11] to overcome high-noise, low-contrast problem in infrared images. A Boolean map is a spatial representation that divides an image into two complementary components: the component that is selected and the component that is not selected. Zhang et al. [12] proposed a saliency detection algorithm based on this concept (namely BMS). However it produced imprecise results when applied to infrared images (Fig. 1).

In the proposed method, by constructing multiple Boolean maps, we exposed image features in different threshold levels, solving the low-contrast problem. Combining with an adaptive thresholding algorithm, we successfully suppressed the noise in infrared images. The complete algorithm is illustrated in Fig. 2 and described in section 2. Section 3 demonstrates the performance of the proposed method and comparison with BMS. This paper is concluded in section 4.

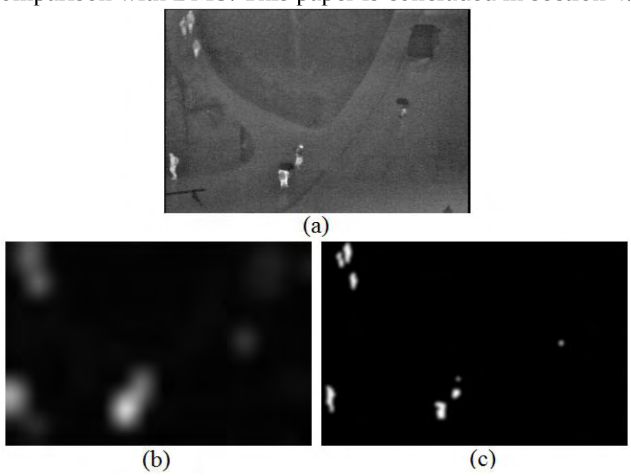

Figure 1. Saliency detection results: (a) input image, (b) result from BMS, (c) result from the proposed method.

2. Proposed method

Given an infrared image I whose pixel values lie in the interval [0, 255], in the first step, $N - 1$ Boolean maps are calculated by

$$B_i = \mathbf{Th}(I, \theta_i),\ i \in \{1, 2, ..., N-1\}$$

where the function $\mathbf{Th}(\cdot)$ assigns 1 to pixels in I whose values are greater than θ_i and 0 otherwise. The value of θ_i is defined as

$$\theta_i = \left\lfloor \frac{i}{N} \times 255 \right\rfloor$$

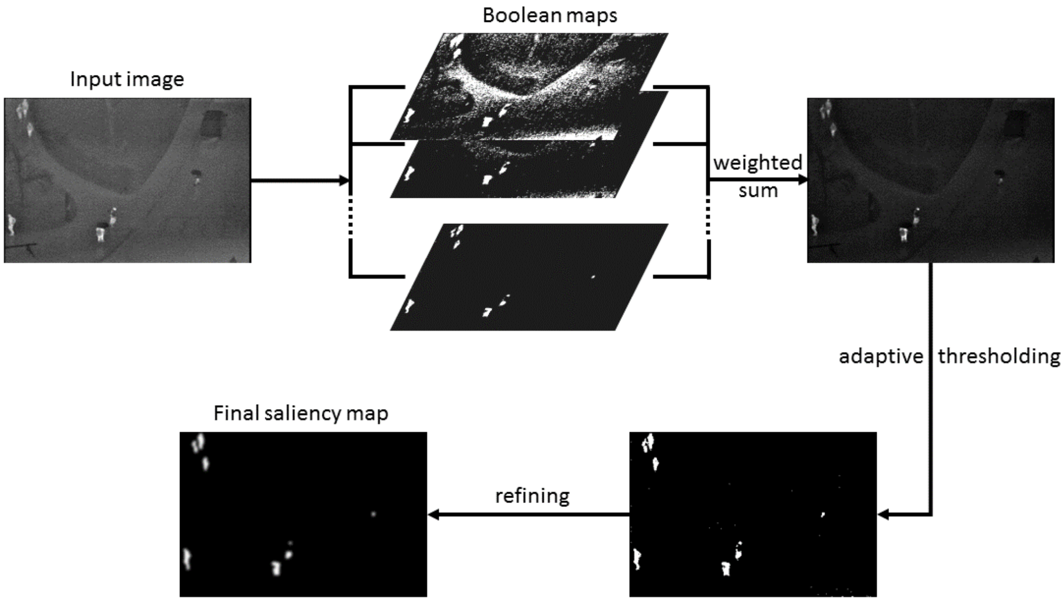

Figure 2. The flow chart of proposed method

We do not perform thresholding at i = 0 (minimum gray level) and $i = N$ (maximum gray level) because the thresholding results in these cases are meaningless, the image will be black or white entirely. After obtaining the Boolean maps, the initial saliency map is calculated as a weighted sum of Boolean maps as follow

$$S = \sum_{i} \omega_i B_i$$

where $\omega_i = \theta_i / 255$, which means Boolean maps that created by higher thresholds have higher weights. This weighting scheme is based on the observation that salient objects in infrared images are usually small and bright. Because the values of pixels in B_i are 0 and 1, before going to the next step, the values of pixels in S are rescaled to the range [0, 255] (rounded down).

The initial saliency map S may contain noise due to the nature of infrared imaging. To eliminate the noise, we threshold the initial map by using a modified version of Otsu's method proposed in [13]. This technique uses entropy information as the additional weighting scheme for Otsu's method. The modified version is better than original method in highlighting small salient regions and not affected by noise. The saliency map is further refined by using opening, a morphological operation. Then we apply a Gaussian filter to create the final saliency map.

3. Experimental results

We evaluated the performance of the proposed method using OSU Thermal Pedestrian Database from [14]. This data set contains infrared image sequences from a camera mounted on rooftop of a building. The main objects of interest in this data set are pedestrians. The images are in gray scale and have a resolution of 360×240 pixels. Because this data set is used for pedestrian detection in the first place, only bounding boxes around pedestrians are provided as the ground truth. We considered pedestrians as salient objects, hence we used the bounding boxes as the indicators for salient regions.

The main parameter in the proposed method is N, the number of thresholding levels. If N is small, the number of exposed features is insufficient. If N is large, some Boolean maps will become redundant because the threshold values are very close to each other. In the experiments, we selected N = 20 for the best results. We also compared our method with BMS [12], a Boolean map-based saliency detection technique. The implementation of BMS was obtained from the author's website.

The saliency detection results are illustrated in Fig. 3. The original infrared images are shown in the left column (Figs. 3a,d,g,j,m), in which the pedestrians are highlighted by yellow bounding boxes. Each row shows the detection results by BMS and the proposed method for the respective infrared image. As can be seen, the proposed method produced better and clearer saliency maps for this data set (Figs. 3c,f,i,l,o). In particular, it captured the general shapes of the pedestrians. The detected salient regions locate accurately inside the bounding boxes. BMS produced unclear salient regions, especially for regions that are close to image boundaries (in Figs. 3e,k,m, the pedestrians at top left corners are unclear). It is worth noting that for BMS, we used maps of eye fixation predictions instead of salient object detection results, because the saliency maps were very imprecise.

4. Conclusions

Saliency detection research plays an important role in several vision-based applications. Saliency detection in infrared images remains a challenging problem due to the nature of input data. In this paper, we introduce a simple approach for saliency detection in infrared images. The input image is thresholded into several Boolean maps, the number of maps is selected based on observation for the best results. Then an initial saliency map is calculated as a weighted sum of created Boolean maps. The final, high-quality saliency map is constructed by further refining the initial map with thresholding, morphology operation, and Gaussian filter. The experiment showed that the proposed method produced high performance results when applied to real-life infrared image sequences.

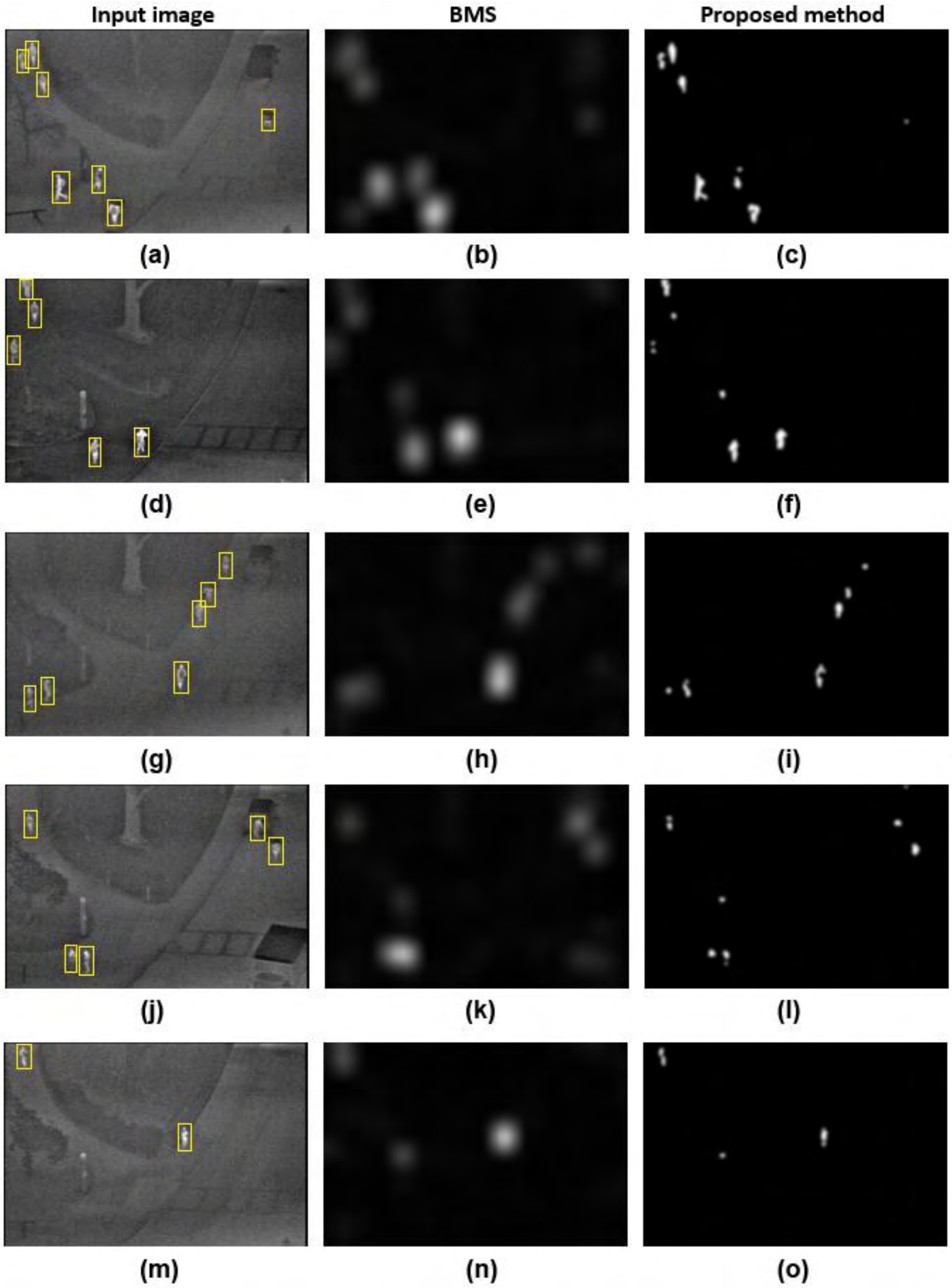

Figure 3. Saliency detection results from BMS and the proposed method. Left column: infrared images with pedestrians inside yellow bounding boxes. Center column: results from BMS. Right column: results from our proposed method.

References

[1] Achanta, R., Susstrunk, S.: Saliency detection for content-aware image resizing. In: 2009 16th IEEE International Conference on Image Processing (ICIP). (2009) 1005–1008

[2] Wulf, S., Zolzer, U.: Visual saliency guided mode decision in video compression based on Laplace distribution of DCT coefficients. In: 2014 IEEE Visual Communications and Image Processing Conference. (2014) 490–493

[3] Mukherjee, P., Lall, B.: Saliency and KAZE features assisted object segmentation. Image and Vision Computing 61 (2017) 82–97

[4] Mesquita, R.G., Mello, C.A.B.: Object recognition using saliency guided searching. Integrated Computer-Aided Engineering 23(4) (2016) 385–400

[5] Borji, A., Cheng, M.M., Jiang, H., Li, J.: Salient object detection: A benchmark. IEEE Transactions on Image Processing 24(12) (2015) 5706–5722

[6] Bylinskii, Z., Judd, T., Borji, A., Itti, L., Durand, F., Oliva, A., Torralba, A.: MIT saliency benchmark

[7] Tsukamoto, T., Esashi, M., Tanaka, S.: Infrared-to-visible transducer using temperature sensitive Eu(TTA)3 on self-suspended thin film for inexpensive thermal imaging device. In: 2013 IEEE 26th International Conference on Micro Electro Mechanical Systems (MEMS). (2013) 421–424

[8] Li, W., Pan, C., Li-xiong, L.: Saliency-based automatic target detection in forward looking infrared images. In: 2009 16th IEEE International Conference on Image Processing (ICIP). (2009) 957–960

[9] Li, L., Zheng, Y., Zhou, F.: Contrast and distribution based saliency detection in infrared images. In: 2015 IEEE 17th International Workshop on Multimedia Signal Processing (MMSP). (2015) 1–6

[10] Qin, S., Wang, L., Cheng, H., Feng, Q., Zhang, M., Gao, C.: Infrared image saliency detection based on human vision and information theory. In: 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). (2016) 484–488

[11] Huang, L., Pashler, H.: A Boolean map theory of visual attention. Psychological Review 114(3) (2007) 599–631

[12] Zhang, J., Sclaroff, S.: Exploiting surroundedness for saliency detection: A Boolean map approach. IEEE Transactions on Pattern Analysis and Machine Intelligence 38(5) (2016) 889–902

[13] Truong, M.T.N., Kim, S.: Automatic image thresholding using Otsu's method and entropy weighting scheme for surface defect detection. Soft Computing (2017)

[14] Davis, J.W., Keck, M.A.: A two-stage template approach to person detection in thermal imagery. In: Seventh IEEE Workshops on Application of Computer Vision, 2005. WACV/MOTIONS'05 Volume 1 (2005) 364–369