Exploring Gradient Information in the Background Subtraction Task

Teera Siriteerakul

King Mongkut's Institution of Technology Ladkrabang,
Bangkok, Thailand 10520, kfteera@kmitl.ac.th

Abstract. Background subtraction is one of the fundamental problems where its solutions can be applied to the other segmentation problems as well. The reported techniques often rely on intensity information to differentiate background pixels from the foreground ones without consider any relation between pixels. To improve the differentiation ability, this paper introduces additional information to be used base on pixel gradient. The analysis given states that the threshold for gradient difference can be selected based on the threshold used for intensity difference. At the end, the experimental results show sufficient improvements of incorporating gradient information in the background differencing task.

Keywords: Background subtraction, Gaussian noise, Image processing, Region of interest

1. Introduction

Background subtraction (a.k.a background differencing) is one of the fundamental problems in Computer Vision. Several vision based systems such as surveillance, object detection, or motion tracking need to detect foreground region before other techniques can be applied.

There are vast differences of pre-condition to the task of background subtraction. Example of the varieties include indoor/outdoor scene, day/night time, or fix/moving camera. Each of these varieties leads to techniques with different approaches. Piccardi [7] and Benezeth et al. [1] provide reviews of the techniques and the commonly-implemented algorithm on this problem.

There are several reported techniques utilizing different kinds of information to differentiate foreground pixels from the background. Jain et al. [4] initiates work in this area by using "large" intensity difference as the indication of foreground status. Since then, there are several reported works using intensity and/or color information to detect foreground pixel [2, 6, 8, 9]. Later, Jabri et al. [3] provided a technique to fuse color and edge information to detect foreground regions. Similarly, Javed et al. [5] offered a way to use gradient information to eliminate false positive pixels (pixels which should be background but judged as foreground). Finally, Yoshida [10] demonstrated that the status of neighbouring pixels can be used as information for the background differencing task.

This paper introduces background differencing technique based on pixel intensity and gradient information. The fundamental idea is relatively simple. Given b(i,j) and t(i,j) be pixels in background and target image, the pixel t(i,j) is judged to be background pixel if and only if the difference in intensity and in gradient of b(i,j) and t(i,j) are below thresholds.

The contribution of this paper is the analysis of using gradient information as complement information to background subtraction task. The fundamental idea is that: If the pixel is not changing, both its color and its gradient should stay the same. On the other hand, this paper also provides an analysis to show that the threshold of the gradient difference can be selected based on the threshold of the intensity difference. The experimental result will show that, by collaborating gradient information, the error pixels can be reduced by 2 percents.

The rest of this paper is organized as follows. The next section will give analysis of the proposed algorithm followed by its outline. After that, two experimental results are shown to demonstrate strength and application of the proposed algorithm. Finally, the concluding discussion will be given in the last section.

2. Background Differencing with Gradient Information

Typically, a fixed-threshold method judge a pixel (as a foreground or background) by comparing a fix number (threshold) to the difference of the intensity level of the pixel from background and the target frame. The label L(i,j) of a pixel t(i,j) can be either B (for background) or F (for foreground) which is given by

$$L(i,j) = \begin{cases} B & \mathrm{e(i,j)} \leq \mathrm{T}, \\ F & \text{otherwise}, \end{cases} \tag{1}$$

where e(i,j)=|t(i,j)-b(i,j)| and T is the given threshold.

In the problem of background differencing, the threshold is given to counter the always-presented image noises. The observed frame difference can be written as e(i,j) = |(t(i,j)+nt)-(b(i,j)+nb)| where nt and nb are mutually independence noises. Therefore, given the image noise modeled using Gaussian distribution with zero mean and the variance σ2, the noise attached to e(i,j) is also Gaussian with the variance 2σ2.

The proposed algorithm handles the additional gradient information analogue to the intensity information. In this paper, the gradient information pG(i,j) of a pixel p(i,j) is given as a 2D vector as [px(i,j),py(i,j)]T. Thus, the difference in gradient, eG(i; j), between b(i,j) and t(i,j) can be calculated using Euclidean distance by

$$e_G(i,j) = \sqrt{(t_x(i,j) - b_x(i,j))^2 + (t_y(i,j) - b_y(i,j))^2}. \tag{2}$$

Since each gradient components in the right hand side of the above equation are often approximated by Sobel operator (see fig. 1), the noises of these components will be Gaussian with variance 8σ2.

$$Sobel_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ +1 & 0 & +1 \end{bmatrix}, Sobel_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}$$

Fig. 1: Sobel Kernels.

To approximate the variance of $e_G(i,j)$, we can set $t_x(i,j)=b_x(i,j)$ and $t_y(i,j)=b_y(i,j)$ the eq. (2) becomes

$$n_G = \sqrt{(n_{tx} + n_{bx})^2 + (n_{ty} + n_{by})^2}, \tag{3}$$

where n_{tx}, n_{bx}, n_{ty}, n_{ty} are noise from the gradient calculation with variance $8\sigma^2$. With simple calculation from above equation, the resulting n_G is Gaussian with variance $(\sqrt{512})\sigma^2$. Given the variance $2\sigma^2$ of the intensity difference, the threshold of gradient difference can be set to be approximately 11 times of the threshold of the intensity difference. The experiments (excluded from this paper) were done to confirm this result.

The decision rule in the eq. (1) is modified to incorporate gradient information as

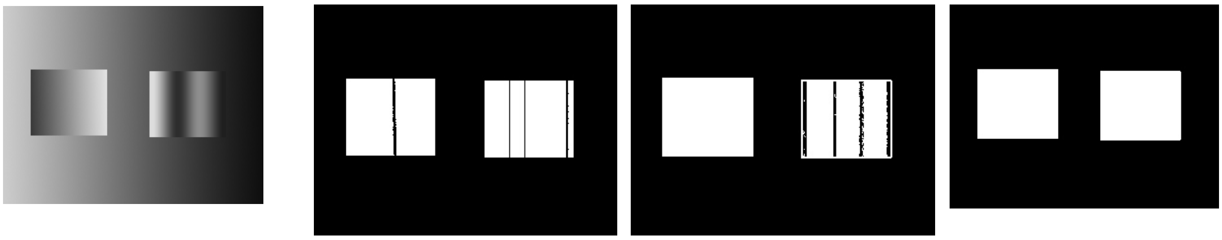
$$L(i,j) = \begin{cases} B & e_I(i,j) \leq T_I \text{ and } e_G(i,j) \leq T_G, \\ F & \text{otherwise}, \end{cases} \tag{4}$$

where the intensity difference e(i,j) is renamed to $e_I(i,j)$, the gradient difference is named $e_G(i,j)$, and the threshold for intensity and gradient are T_I and T_G accordingly.

With the earlier analysis result, the gradient threshold can be set to be 11 times of the intensity threshold. Thus, the additional differentiation ability of gradient information comes without the burden of having to fine-tuning the gradient threshold.

3. Experimental Results

There are to set of experiments to demonstrate and evaluate the performance of the additional gradient information. First, the proposed method is performed on a set of images designed to demonstrate the strength of the algorithm. Then, the method is compared with a fixed-threshold method on images taken from a notebook camera.

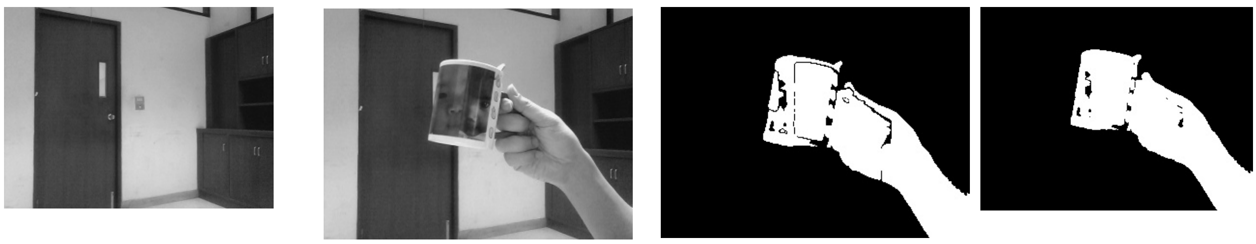


(a) (b) (c) (d)

Fig. 2 (a) target image, (b) intensity threshold, (c) gradient threshold, and (d) combination of intensity and gradient threshold.

3.1. Demonstration on predefined images

In this experiment, the proposed algorithm is performed on a set of simple images. As demonstrated in Fig. 2(a), the left area of the image contains a region which having similar intensity as its background, while all pixels in the region has opposite gradient direction. On the other hand, the right area contains some parts where their pixels intensities are the same as background pixel. Within these areas with similar intensity as background, some regions have opposite gradient direction while some regions have the same gradient direction as background but with stronger magnitude. Also, there are some regions in the right area where their gradient are similar in both direction and magnitude with the background.



(a) (b) (c) (d)

Fig. 3 (a) background image, (b) target image, (c) intensity threshold, and (d) combination of intensity and gradient threshold.

3.2. Demonstration on practical images

The results of background differencing techniques, proposed and fixed-threshold, are shown in Fig. 3. The combination result illustrated in Fig. 3(d) shows noticeable improvement over the intensity-threshold result illustrated in Fig. 3(c). Furthermore, when define percent of error as 100*[number of incorrectly labeled pixels]/[number of total foreground pixels], the proposed method reduces the error by roughly 2% in this case.

4. Conclusion

This paper gives analysis as well as empirical evidents of how gradient information can be useful in the task of background differencing. The analysis given allows the gradient information to be used in combine with intensity information without having to fine-tune a separate gradient threshold. In addition, the analysis given by this paper should be useful for any task requires pixels differentiation.

5. References

[1] Y. Benezeth; P.M. Jodoin; B. Emile; H. Laurent; C. Rosenberger. Review and evaluation of commonly-implemented background subtraction algorithms. *Proceeding of the 19th International Conference on Pattern Recognition,* 2008, pp. 1–4.

[2] I. Haritaoglu; D. Harwood; L.S. Davis. W4: real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Learning,* 2000, 22, pp. 809–830.

[3] S. Jabri; Z. Duric; H. Wechsler; A. Rosenfeld. Detection and location of people in video images using adaptive fusion of color and edge information. *Proceeding of the 15th International Conference on Pattern Recognition,* 2000, 4, pp. 627–630.

[4] R. Jain; D. Militzer; H. Nagel. Separating nonstationary from stationary scene components in a sequence of real world tv-images. *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, 1977, 2, pp. 612–618.

[5] O. Javed; K. Shafique; M. Shah. A hierarchical approach to robust background subtraction using color and gradient information. *Proceeding of the Workshop on Motion and Video Computing*, 2002, pp. 22–27.

[6] N.M. Oliver; B. Rosario; A.P. Pentland. A Bayesian computer vision system for modeling human interactions. *IEEE Transactions on Pattern Analysis and Machine Learning*, 2000, 22, pp. 831–843.

[7] M. Piccardi. Background subtraction techniques: a review. *International Conference on Systems, Man and Cybernetics*, 2004, 4, pp. 3099–3104.

[8] C. Stauffer; W.E.L. Grimson. Learning patterns of activity using real-time tracking. *IEEE Transactions on Pattern Analysis and Machine Learning*, 2000, 22, pp. 747–757.

[9] C. Wren; A. Azarbayejani; T. Darrell; A. Pentland. Pfinder: real-time tracking of the human body. *In Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, 1996, pp. 51–56.

[10] Y. Yoshida. Background differencing technique for image segmentation based on the status of reference pixels. *Proceeding of the International Conference on Image Processing*, 2004, 5, pp. 3487–3490.