Background Subtraction using Random Walks with Restart

Tae Hoon Kim, Kyoung Mu Lee and Sang Uk Lee

Dept. of EECS, ASRI
Seoul National University
Seoul, Korea
E-mail: thkim@diehard.snu.ac.kr, kyoungmu@snu.ac.kr, sanguk@ipl.snu.ac.kr

ABSTRACT

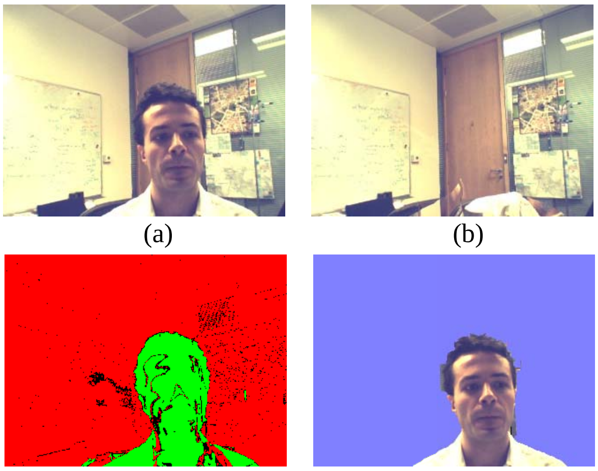
Automatic segmentation of foreground from background in video sequences has attracted lots of attention in computer vision. This paper proposes a novel framework for the background subtraction that the foreground is segmented from the background by directly subtracting a background image from each frame. Most previous works focus on the extraction of more reliable seeds with threshold, because the errors are occurred by noise, weak color difference and so on. Our method has good segmentations from the approximate seeds by using the Random Walks with Restart (RWR). Experimental results with live videos demonstrate the relevance and accuracy of our algorithm.

Keywords: background subtraction, random walks with restart

1. Introduction

Automatic segmentation of foreground is an important issue in computer vision. In most recent works, two different segmentation methods were proposed. The first one is bilayer segmentation where a complex energy function depending on both segmentation labels and training image data is setup such as Criminisi et al [2]. The segmentation is then estimated by minimizing this energy function. The energy encodes terms that enforce coherence of segmentation in space-time as well as with image data. Since the segmentation results are changed according to the training stage where some hand-labeling of a few sequences is necessary to build color and motion models, it is important to find the proper models. The second method is background subtraction such as Sun et al [6]. If a static background image is given or generated, it can be automatically resulted by comparing |frame-background| with a pre-defined threshold *T*. If the difference of a pixel is larger than *T*, we then classify it as foreground. Although it is easily applied, it is difficult to deal with some errors by noise, less color difference and so on. In this paper, we assume that the foreground is moving while the background remains static. Therefore, we focus on the background subtraction of second type.

We propose a new background subtraction algorithm based on the Random Walks with Restart (RWR) [5][7][4] that can solve the segmentation problem from approximate seeds effectively. Fig. 1 shows the overall process of our



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

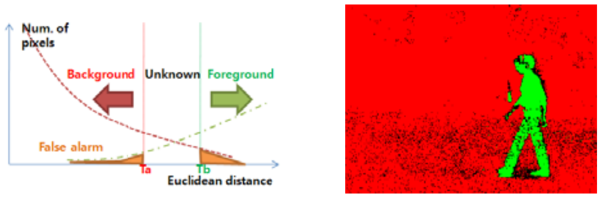
Fig. 1: An example of our background subtraction model. (a) One frame in video. (b) Median image generated from a partial input video as a background image. (c) Initial seeds with the pre-defined thresholds. (d) Our segmentation result based on initial seeds.

method from the initial seeds with thresholds to the resulting foreground segmentation. We use the test videos where the foregrounds are moving around in the static background. If the background image is not given, it can be automatically estimated by taking the median of a relatively small subset of the frames in Fig. 1(b). After computing the initial seeds by comparing |Fig. 1(a) – Fig. 1(b)| with the pre-defined thresholds in Fig. 1(c), the resulting foreground segmentation can be obtained in Fig. 1(d). Our algorithm has some advantages. First, it has low complexity, because any color or motion models are not used. Also, we use the soft constraint that the resulting labels of initial seeds can be changed according to the labels of the neighboring pixels by the RWR. It gives the efficiency that reliable segmentations can be obtained from the inaccurate initial seeds. Finally, our segmentation results have good performance, compared with Graph Cuts (GC) [1] and Random Walks (RW) [3]. It was proved that the RWR is the state of the art method for seeded image segmentation in [4].

The paper is organized as follows. In Section 2, we introduce our proposed background subtraction algorithm and explain that algorithm in detail. The experimental results are shown in Section 3. Finally, we discuss our approach and give conclusions in Section 4.

(a) (b)

Fig. 2: An example of our estimated background image. (a) Input frames in video. (b) Resulting background image by median filtering of some frames.



(a) (b)

Fig. 3: An example of our initial seeds with thresholds Ta , Tb . (a) A detailed explanation. (b) Initial seeds (red: background, green: foreground, black unknown).

2. Proposed Algorithm

Now, we describe the process of our proposed foreground segmentation algorithm in detail.

2.1 Estimate a background image

If the foregrounds are moving around in the static background, we can estimate an approximate background image by taking the median of a relatively small subset of the frame. Given a few video frames in Fig. 2(a), the resulting background image is estimated in Fig. 2(b). It is an approximate background, not exact.

2.2 Find the initial seeds

Given an image, the initial seeds are estimated as shown in Fig. 3(a). After the difference image, |frame-background|, is obtained, the pixels that have larger value than a threshold Tb in this difference image are estimated as foreground seeds in Fig. 3(b). Also, the pixels that have smaller value than a threshold Tb in this difference image are estimated as background seeds. Since the estimated background image is not exact solution, the labels of seeds are not reliable. Therefore, our algorithm uses the soft constraint that each seed tends to preserve the initial label.

2.3 Background Subtraction

Let us consider the foreground segmentation as a labeling problem in which each pixel $x^i \in X = \{x^1,...,x^N\}$ is to be assigned a label $l \in L = \{foreground, background\}$. In a generative approach, the posterior probability is obtained using Bayesian rules:

$$p(l \mid x^i) = \frac{p(x^i \mid l)p(l)}{\sum_l p(x^i \mid l)p(l)}, \tag{1}$$

where the sum in the denominator is taken over all labels. The likelihood $p(x^i \mid l)$ can be estimated by

$$\begin{aligned} p(x^i \mid l) &= \sum_{m=1}^{M_l} p(x^i \mid x^{m,l}, l)p(x^{m,l} \mid l) \\ &\approx \sum_{m=1}^{M_l} p(x^i \mid x^{m,l}, l) \end{aligned}, \tag{2}$$

where $X^l = \{x^{1,l},...,x^{M_l,l}\}$ is a set of the M_l seeds with label l and the seed distribution $p(x^{m,l} \mid l)$ is defined by a uniform distribution. The likelihood of each pixel is modeled by a mixture of distribution from each seed.

The process of our proposed background subtraction method is briefly described as follows. First, we construct a weighted graph. Then, we define $p(x^i \mid x^{m,l}, l)$ as the steady-state probability that a random walker starting at each seed $x^{m,l}$ stays at a pixel x^i in this graph. After computing these steady-state probabilities using RWR, we can obtain the likelihood $p(x^i \mid l)$ using (2) and, finally assign the label with maximum posterior probability in (1) to each pixel.

2.3.1 Graph Model

Given an image I, let us construct an undirected graph $G = (V, E)$ with nodes $v \in V$, and edges $e \in E$. Each node v^i uniquely identifies an image pixel x^i. The edge e^{ij} between two nodes v^i, v^j is determined by the neighborhood system. The weight $w^{ij} \in W$ is assigned to the edge e^{ij}. It measures the likelihood that two neighboring nodes have the same label. In this work, it is defined as the Gaussian weighting function [1][3][4] in the 4 neighborhood system as follows.

$$w^{ij} = \exp\left(-\frac{|g(x^i) - g(x^j)|^2}{\sigma}\right) \tag{3}$$

where the function $g(\cdot)$ indicates the image colors in *Lab* color space.

2.3.2 Foreground Segmentation

Suppose a random walker starts from a *m*-th seed $x^{m,l}$ of label l in this graph G. The random walker iteratively transmits to its neighborhood with the probability that is proportional to the edge weight between them. Also at each step, it has a restarting probability c to return to the seed $x^{m,l}$. After convergence, we obtain the steady-state probability $r^{(i,m),l}$ that the random walker will finally stay at a pixel x^i. In this work, we use this steady-state probability as the distribution $p(x^i \mid x^{m,l}, l)$ in (2) such as

$$p(x^i \mid x^{m,l}, l) \approx r^{(i,m),l}. \tag{4}$$

By denoting $r^{(i,m),l}$, $i = 1,...,N$ in terms of a vector $\vec{r}^{m,l}$ and defining an adjacency matrix $W = [w^{ij}]_{N \times N}$, RWR can be formulated as follows [5][7][4].

$$\begin{aligned} \vec{r}^{m,l} &= (1-c)P\vec{r}^{m,l} + c\vec{b}^{m,l} \\ &= c(I - (1-c)P)^{-1}\vec{b}^{m,l} \\ &= Q\vec{b}^{m,l} \end{aligned} \tag{5}$$

where $\vec{b}^{m,l} = [b^i]_{N \times 1}$ is the indicating vector with $b^i = 1$ if $x^i = x^{m,l}$ and $b^i = 0$ otherwise, and the transition matrix $P = [p^{ij}]_{N \times N}$ is the adjacency matrix W row normalized:

(a) Initial seeds (b) GC (c) RW (d) Our Algorithm

Fig. 4: Comparison of our algorithm with GC and RW on 'A_anto' sequence (200 frames) [8]. (b),(c) and (d) are the segmentation results of GC, RW and our algorithm, respectively. ($c = 0.0002$)

$$P = D^{-1}W \quad (6)$$

where $D = diag(d^1,...,d^N)$, $d^i = \sum_{j=1}^{N} w^{ij}$. If $\vec{r}^{m,l}$ are inserted into (2) by our definition, the likelihoods $p(x^i \mid l)$ are achieved such as:

$$[p(x^i \mid l)]_{N\times 1} = Q\vec{d}^l \quad (7)$$

where $\vec{d}^l = [d^i]_{N\times 1}$ is the vector with $d^i = 1$ if $x^i \in X^l$ and $d^i = 0$ otherwise.

Assume that the prior probability $p(l)$ in (1) is uniform. Using this likelihood in (7), the decision rule of each pixel x^i for segmentation is as follows:

$$R^i = \arg\max_l p(l \mid x^i) = \arg\max_l p(x^i \mid l) \quad (8)$$

The segmentation is obtained by assigning the label R^i to each pixel x^i.

3. Experimental Results

Our algorithm has two parameters: a color variance σ and a restarting probability c. They are fixed with the same value for all the segmentation algorithms we tested. We compare the performance of our algorithm with GC [1] and RW [3] on live videos. We utilized a dataset of live videos [8][9]. In Fig. 4 and Fig. 5, the segmentations were produced from the three different algorithms on these videos. The foreground boundary is drawn in red color overlaid on the original frame. Compared with the segmentations from GC and RW, our algorithm has better segmentations qualitatively. GC and RW have a hard constraint that the seeds hold the initial labels. Thus, the selection of the initial seeds is very important. To improve performance, the additional constraint is needed such as [6]. Our algorithm is robust to the initial errors in Fig. 4(d). This comparison confirms the relevance and accuracy of our algorithm.

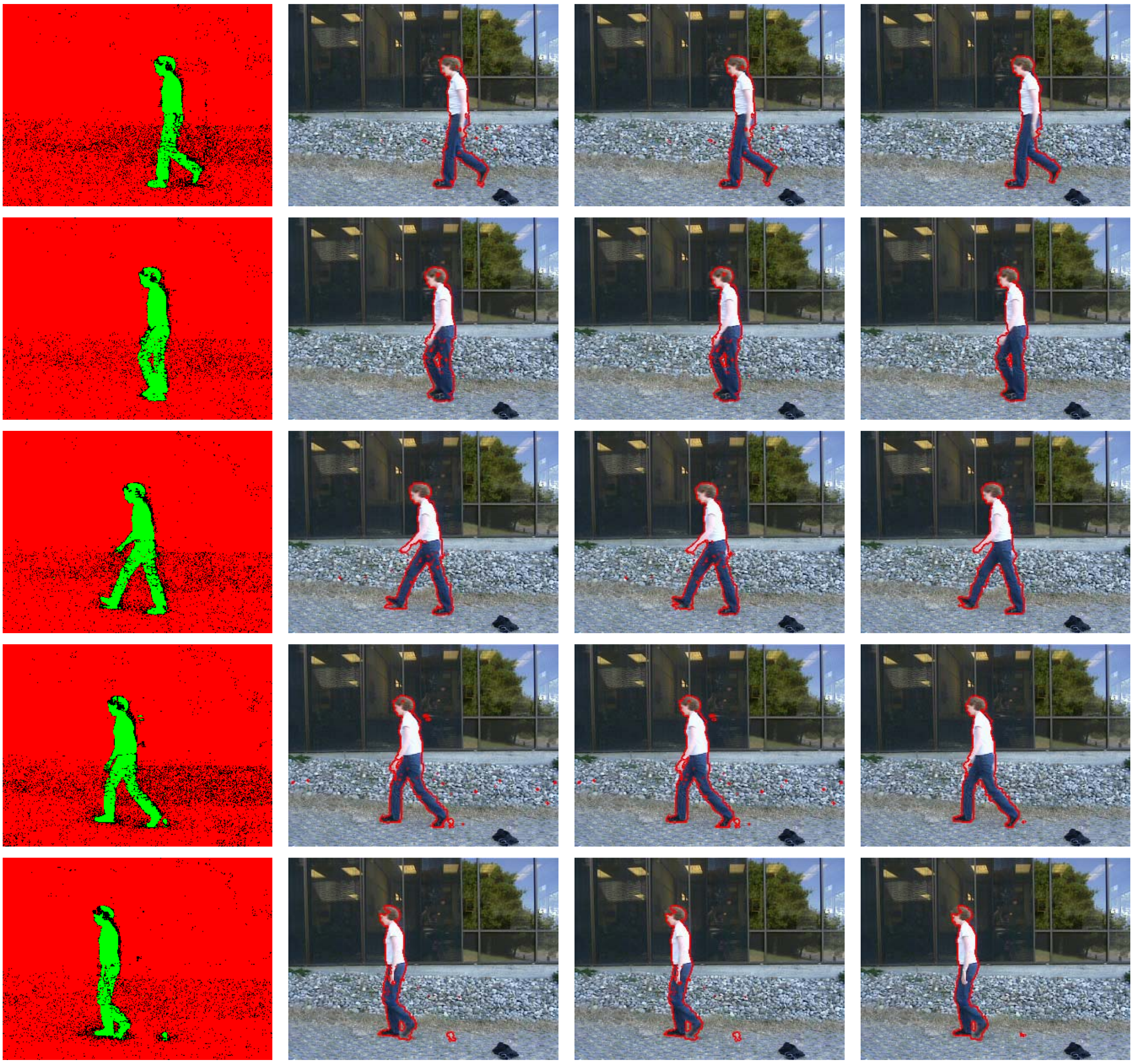
4. Conclusion

This paper presents a novel generative background subtraction algorithm in the Bayesian Framework. By using RWR, our work produces significant improvement in performance as shown in the experiment.

For the computation of RWR, the restarting probability c was chosen empirically. However it is not optimal for every video. If we can control it well, better segmentations will be obtained. Thus our future work will include the automatic selection of the optimal value of this parameter.

5. Acknowledgement

This research was supported in part by the Defense Acquisition Program Administration and Agency for Defense Development, Korea, through the Image Information Research Center at Korea Advanced Institute of Science & Technology under the contract UD070007AD, and in part by the IT R&D program of MKE/IITA. [2008-F-030-01, Development of Full 3D Reconstruction Technology for Broadcasting Communication Fusion]



(a) Initial seeds (b) GC (c) RW (d) Our Algorithm

Fig. 5: Comparison of our algorithm with GC and RW on 'human_in' sequence (30 frames) [9]. (b),(c) and (d) are the segmentation results of GC, RW and our algorithm, respectively. ($c = 0.004$)

6. REFERENCES

[1] Y. Boykov and G. Funka-Lea, "Graph Cuts and Efficient N-D Image Segmentation," *International Journal of Computer Vision*, 2006

[2] A. Criminisi, J. Shotton, A. Blake, C Rother and PHS Torr, "Bilayer Segmentation of Live Video," *In Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2006.

[3] L. Grady, "Random Walks for Image Segmentation," *Pattern Analysis and Machine Intelligence*, 2006

[4] T. H. Kim, K. M. Lee and S. U. Lee, "Generative Image Segmentation using Random Walks with Restart," *In Proc. European Conference on Computer Vision*, 2008.

[5] J. Y. Pan, H. J. Yang, C. Faloutsos and P. Duygulu, "Automatic multimedia cross-modal correlation discovery," *In Proc. Knowledge Discovery & Data Mining*, 2004

[6] J. Sun, W. Zhang, X. Tang and H-Y Shum, "Background Cut," *In Proc. European Conference on Computer Vision*, 2006.

[7] H. Tong, C. Faloutsos and J. Y. Pan, "Fast Random Walk with Restart and Its Applications," *In Proc. International Conference on Data Mining*, 2006.

[8] http://research.microsoft.com/vision/cambridge/i2i/default.htm

[9] http://www.robots.ox.ac.uk/~vgg/data/data-various.html