Semi-Supervised Video Object Segmentation Using Multiple Random Walkers

Won-Dong Jang
http://mcl.korea.ac.kr/~dotol1216

Chang-Su Kim
http://mcl.korea.ac.kr

School of Electrical Engineering
Korea University
Seoul, Korea

In this work, we propose a semi-supervised video object segmentation algorithm, which require user annotations about a desired object at the first frame. Figure 1 is an overview of the proposed algorithm. For each frame, we estimate initial distributions and simulate multiple random walkers (MRW) [1]. We perform these two-step processes from the second to the last frames sequentially to yield a segment track.

First, we estimate initial distributions of the foreground and the background based on the segmentation results of previous frames. To this end, we minimize an energy function, which consists of three terms: color Markov energy, motion Markov energy, and guidance energy. Second, we simulate MRW using the two initial distributions. The movements of the foreground agent $\mathbf{p}_f$ and the background agent $\mathbf{p}_b$ are modeled by

$$\mathbf{p}_f^{(\theta+1)} = (1-\varepsilon)\mathbf{A}_c\mathbf{p}_f^{(\theta)} + \varepsilon\mathbf{r}_f^{(\theta)}, \qquad (1)$$

$$\mathbf{p}_b^{(\theta+1)} = (1-\varepsilon)\mathbf{A}_c\mathbf{p}_b^{(\theta)} + \varepsilon\mathbf{r}_b^{(\theta)}, \qquad (2)$$

where $\mathbf{r}_f$ and $\mathbf{r}_b$ are the restart distributions. $\mathbf{A}_c$ is a transition matrix. With probability $1-\varepsilon$, the agents move on the graph according to the transition matrix $\mathbf{A}_c$. On the other hand, with probability ε, the foreground and background agents are forced to restart with the distributions $\mathbf{r}_f$ and $\mathbf{r}_b$, respectively. We use a restart rule, which is a hybrid of inference and interactive restart rules. The inference restart rule is time-invariant and inferred from the previous segmentation labels. The time-variant interactive restart rule encourages repulsive interactions between the agents.

Experimental results demonstrate that the proposed algorithm outperforms the state-of-the-art conventional algorithms on the SegTrack v2 dataset [2]. To summarize, this paper has three main contributions:

- Development of an effective restart rule for MRW that yields spatially precise and temporally consistent segment tracks.
- Fixation of parameters, which ensures segmentation qualities on general videos.
- Remarkable performance achievement on the SegTrack v2 dataset.

[1] C. Lee, W.-D. Jang, J.-Y. Sim, and C.-S. Kim. Multiple random walkers and their application to image cosegmentation. In *CVPR*, pages 3837–3845, 2015.

[2] F. Li, T. Kim, A. Humayun, D. Tsai, and J. M. Rehg. Video segmentation by tracking many figure-ground segments. In *ICCV*, pages 2192–2199, 2013.
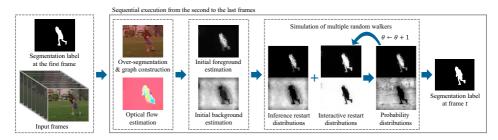
Figure 1: An overview of the proposed algorithm. Using segmentation labels at previous frames, we initialize foreground and background distributions for each frame. Then, we simulate MRW using the inference restart rule and the interactive restart rule. For the segmentation, we compare the foreground and background probabilities at each superpixel.