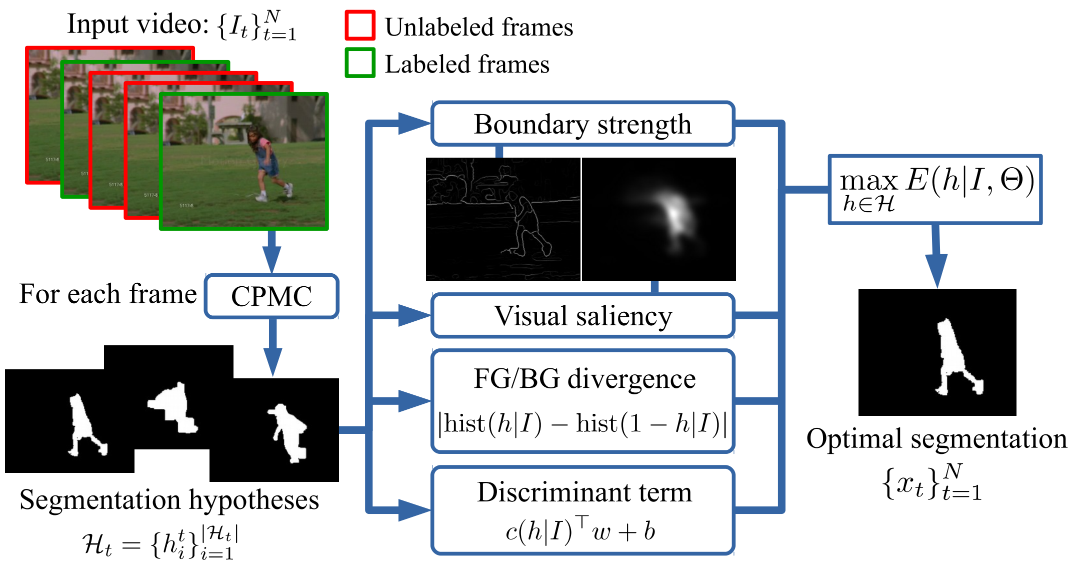
Transductive Video Co-segmentation on The Temporal Trees

Zhihui Fu, Botao Wang, Hongkai Xiong
Image, Video, and Multimedia Communication Lab (IVM Lab)
http://ivm.sjtu.edu.cn
Shanghai Jiao Tong University

ABSTRACT

This paper proposes a novel multi-component video co-segmentation approach to simultaneously separate the foreground from the background in the video frames. To capture the variance of appearance of the foreground object, a multi-component foreground model is developed. Each component of the model characterizes a specific viewpoint/pose/appearance of the foreground object. To learn the parameters of the multi-component model, a transductive learning algorithm is leveraged to "transfer" the information of the labeled frames to the unlabeled frames in a tree-structured model, namely, temporal tree. Each branch of the temporal tree consists of the exemplars of a foreground component, and a transductive support vector regressor is capable of being trained. Experiments show that the proposed method outperforms quite a few state-of-the-art video segmentation algorithms in public benchmark.

Temporal Coherent Video Co-segmentation



Input video: $\{I_t\}_{t=1}^{N}$

🟥 Unlabeled frames

🟩 Labeled frames

For each frame

CPMC

Segmentation hypotheses

$\mathcal{H}_t = \{h_i^t\}_{i=1}^{|\mathcal{H}_t|}$

Boundary strength

Visual saliency

FG/BG divergence

$|\mathrm{hist}(h|I) - \mathrm{hist}(1-h|I)|$

Discriminant term

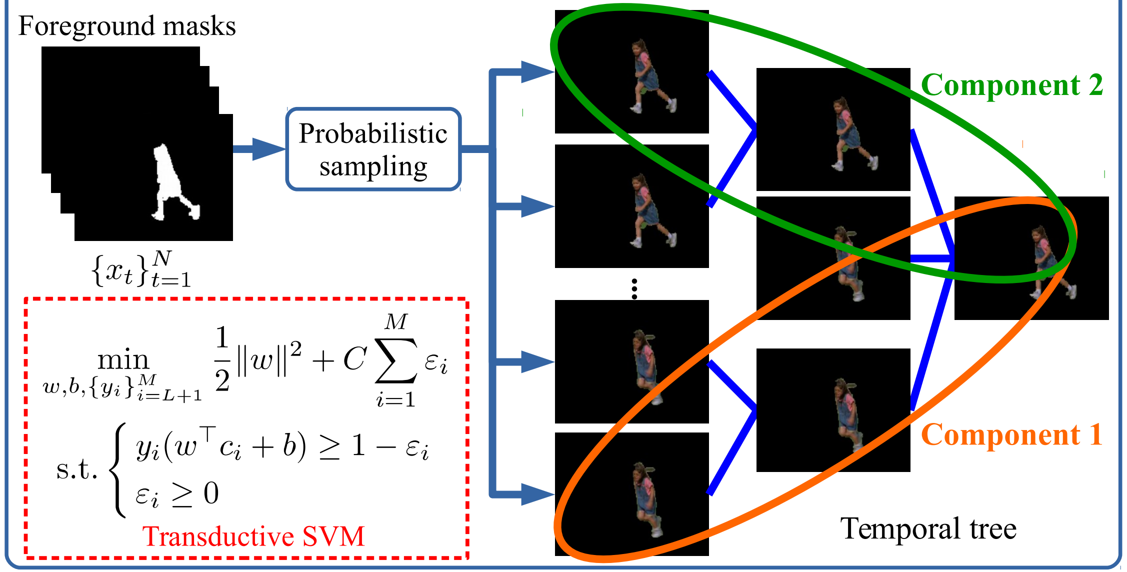
$c(h|I)^\top w + b$

$\max_{h\in\mathcal{H}} E(h|I,\Theta)$

Optimal segmentation

$\{x_t\}_{t=1}^{N}$

Transductive Learning of Multi-Component Foreground Model on the Temporal Tree



Foreground masks

$\{x_t\}_{t=1}^{N}$

Probabilistic sampling

$$\min_{w,b,\{y_i\}_{i=L+1}^{M}} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}\varepsilon_i$$

$$\text{s.t.} \begin{cases} y_i(w^\top c_i + b) \geq 1 - \varepsilon_i \\ \varepsilon_i \geq 0 \end{cases}$$

Transductive SVM

Component 2

Component 1

Temporal tree

Experimental Results

	birdfall	*cheetah*	*girl*	*monkeydog*	*parachute*	*penguin*
Proposed	**190**	**753**	1871	722	387	4841
Joulin et al., *CVPR'12*	988	3279	5321	1125	3245	8932
Chockalingam et al., *ICCV'09*	454	1217	**1755**	683	502	6627
Grundman et al., *CVPR'10*	305	1219	5777	**493**	1202	**2116**
Lee et al., *ICCV'11*	288	905	1785	**493**	**201**	136285