

Causal Video Object Segmentation From Persistence of Occlusions

Brian Taylor Vasiliy Karasev Stefano Soatto
University of California, Los Angeles
<http://vision.ucla.edu/cvos/>

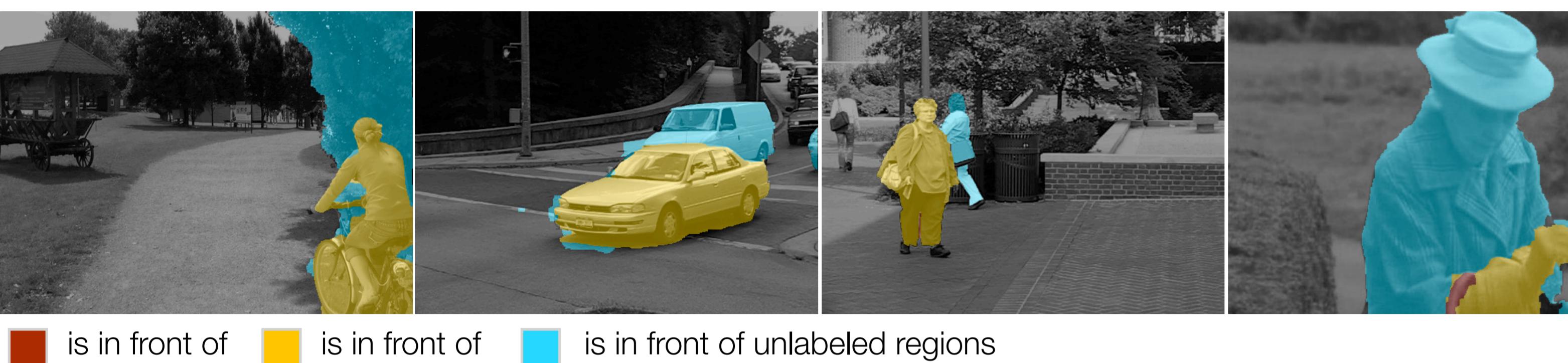
VISIONLAB

our approach

Leverage **occlusion cues** to segment frames into regions ordered by distance from the camera (depth layers) to segment **objects** in video

what is an object? “*a layout of surfaces completely surrounded by the medium*

why occlusions? “[They] are significant in the perception of the layout of the scene as they yield to the relative position of surfaces...” - J.J.Gibson



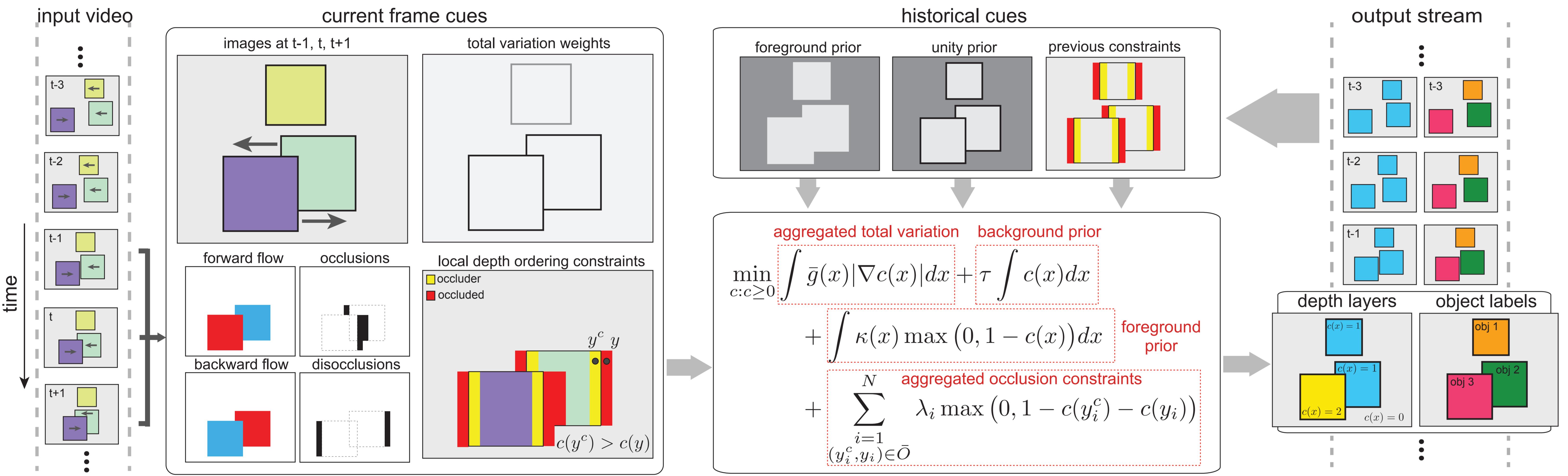
related video segmentation work

- no concept of objects
(oversegmentation)
 - batch, noncausal processing
(scales poorly with video length)
 - restricted to a single moving object
(binary segmentation)
 - tracking?
(requires manual initialization)
 - joint motion+segmentation+layers?
(nonconvex, hard to optimize)

contributions

- process video *causally*
 - design priors to ensure temporally consistent segmentations
 - provide depth-order relationships of objects in the scene wrt viewer
 - introduce an effective scheme for determining occlusion relations
 - segment each frame by solving an efficient convex program

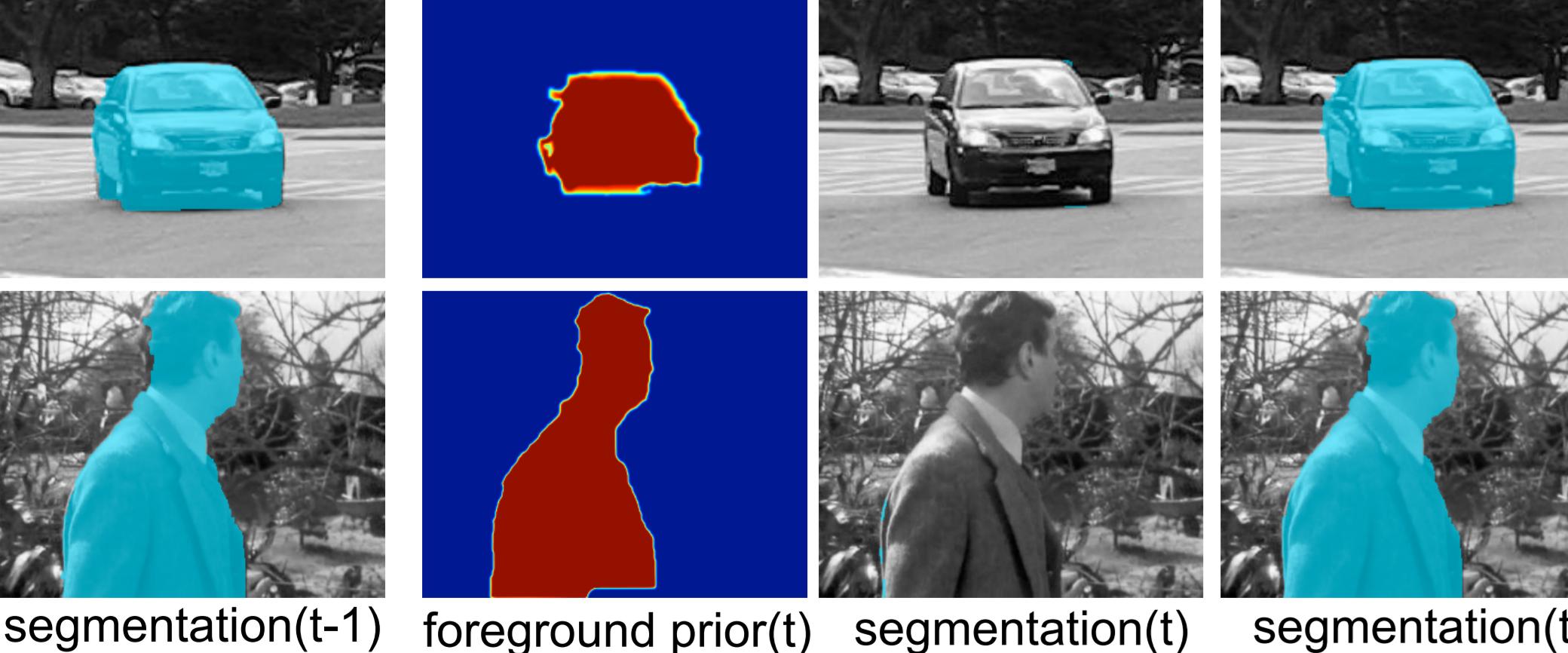
framework overview



historical cues

foreground prior

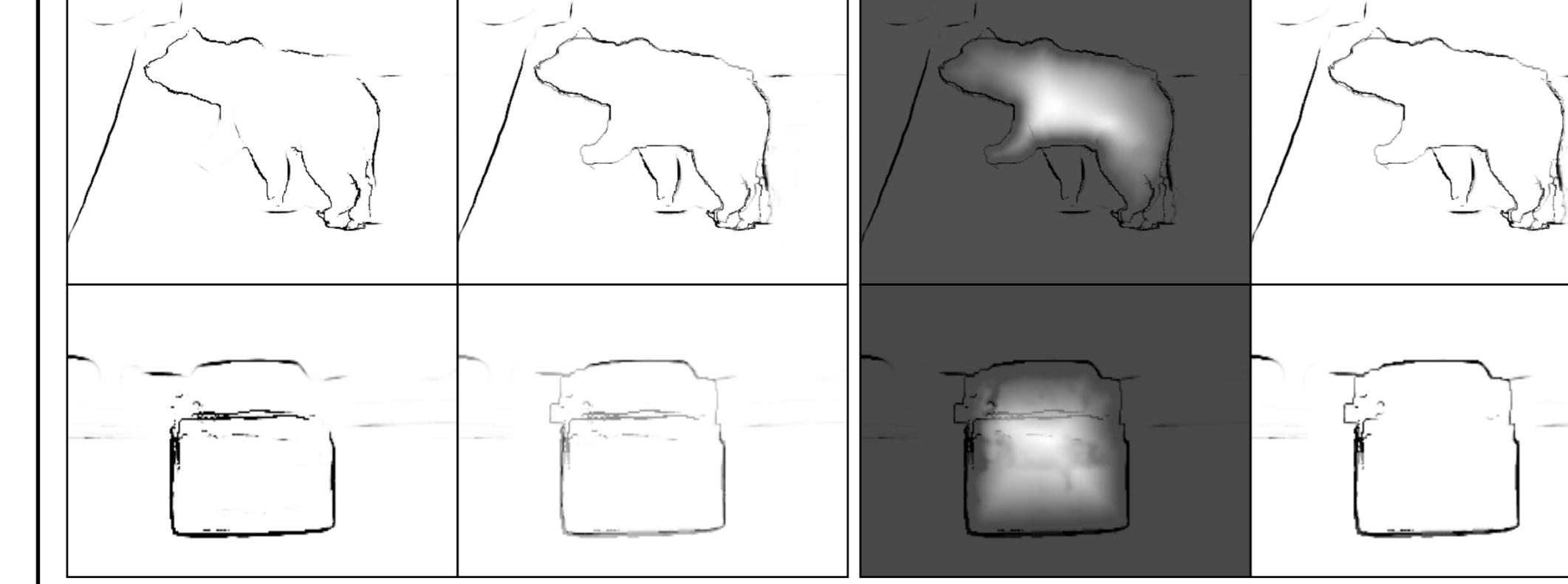
once an object, always an object



- encourages $c(x) \geq 1$ in regions previously assigned to any foreground depth layer

layer unity prio

layer boundaries persist over time



- uses previous segmentations to improve object boundaries
- leverages persistence of object boundaries

aggregated occlusion constraints

retain strong occlusion cues



- combines past occlusion cues with current ones
- leverages persistence of occlusion cues

cues from the past and cues from the current frame are adaptively combined based on the motion in the current frame

