Causal Video Object Segmentation From Persistence of Occlusions

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

**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.Gibson.

**related video segmentation work**

- no concept of objects
- batch, noncausal processing
- restricted to a single moving object
- tracking?
- joint motion+segmentation+layers?

(oversegmentation)
(scales poorly with video length)
(binary segmentation)
(requires manual initialization)
(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**

- input video
- current frame cues
- total variation weights
- local depth ordering constraints
- aggregated total variation
- background prior
- aggregated occlusion constraints
- depth layers
- object labels
- output stream

**historical cues**

- once an object, always an object
- layer unity prior
- aggregated occlusion constraints

- layer boundaries persist over time
- retain strong occlusion cues

- encourages \( c(x) \geq 1 \) in regions previously assigned to any foreground depth layer
- uses previous segmentations to improve object boundaries
- leverages persistence of object boundaries

**results**

- original image
- ground truth
- Grauman'11
- Papazoglou'13
- Grundmann'10
- Brox'14
- ours

**Supported on NSF RI-1422669, ONR N00014-13-1-034, FA8650-11-1-7156**